(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 700 995 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
18.07.2001 Bulletin 2001/29

(51) Int Cl.[7]: **C12N 15/09**, C12N 15/66

(21) Application number: **95306235.3**

(22) Date of filing: **06.09.1995**

(54) Process for production of protein

Verfahren für die Herstellung von Proteinen

Procédé pour la production de protéines

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB IE IT LI NL SE

(30) Priority: **07.09.1994 JP 23859594**
07.11.1994 JP 29602894

(43) Date of publication of application:
13.03.1996 Bulletin 1996/11

(73) Proprietor: **SUNTORY LIMITED**
Kita-ku, Osaka-shi, Osaka (JP)

(72) Inventors:
- **Yabuta, Masayuki**
  Tatebayashi-shi, Gunma (JP)
- **Ohsuye, Kazuhiro**
  Ohta-shi, Gunma (JP)

(74) Representative: **Stoner, Gerard Patrick et al**
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

(56) References cited:
- **APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 44, 1995, pages 118-125, XP002008703 M. YABUTA ET AL.: "Hyperproduction of a recombinant fusion protein of S. aureus V8 protease in E. coli and its processing by OmpT protease to release an active V8 protease derivative"**
- **APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 37, 1992, pages 621-625, XP002008704 P.G. SEEBOTH ET AL.: "In-vitro cleavage of a fusion protein bound to cellulose using the soluble yscF (KEX2) variant"**
- **ENZYME MICROB. TECHNOL., vol. 15, 1993, pages 593-600, XP002008705 I. SUOMINEN ET AL.: "Enhanced revocery and purification of Aspergillus glucoamylase from S. cerevisiae by the addition of poly(aspartic acid) tails"**
- **EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 170, 1987, pages 241-246, XP002008706 C.J. VAN DEN BERGH ET AL.: "Secretion of biologically active porcine prophospholipase A2 ba S. cerevisiae"**
- **BIOTECHNOLOGY, vol. 9, 1991, pages 378-381, XP002008707 R. CONTRERAS ET AL.: "Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by A. nidulans and secretion of mature interleukin-6"**
- **MOL. GEN. GENET., vol. 233, 1992, pages 42-48, XP002008708 C. HANKE ET AL.: "Processing by OmpT of fusion proteins carrying the HlyA transport signal during secretion by the E. coli hemolysin transport system"**

EP 0 700 995 B1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The present invention relates to a process for production of a desired polypeptide using gene recombination techniques. Preferably the desired polypeptides are physiologically active polypeptides, for examples enzymes such as proteases. For example, the present invention relates a process for production of derivatives of Staphylococcus aureus V8 protease.

[0002] As a more specific example, the present invention provides a process for production of an active V8 protease derivative by expressing an insoluble fusion protein of a V8 protease derivative of Staphylococcus aureus origin using an E. coli expressing system, excising the V8 protease derivative from the fusion protein with an ompT protease intrinsic in E. coli in the presence of a denaturating agent and if necessary refolding the V8 protease.

[0003] Staphylococcus aureus (S. aureus) V8 protease is one of the proteases secreted into a culture medium by S. aureus V8 strain. This enzyme was isolated and purified by Drapeau, G. R et al. in 1972, as one of the serine proteases, which is secreted into a culture medium of S. aureus V8 strain and specifically cleaves the C-terminal of glutamic acid and aspartic acid (Jean Houmard and Gabriel R. Drapeu (1972), Proc Natl. Acad. Sci. USA, 69, 3506 - 3509). A DNA nucleotide sequence of the enzyme was determined by Cynthia Carmona et al., in 1987 (Cynthia Carmona and Gregory L. Gray (1987), Nucleic Acids R s. 15, 6757).

[0004] It is believed that the present enzyme is expressed as a precursor having 336 amino acid residues, and secreted as a mature protein by deletion of a prepro sequence of 68 amino acid residues from the N-terminal of the precursor. In addition, it is known that the present enzyme has a repeat sequence of proline-aspartic acid-asparagine at the C-terminal region (amino acid numbers 221 - 256). It is not clear whether or not this repeat sequence is essential for enzymatic activity, and Gray et al. consider that the repeat sequence might function when the enzyme exists as an inactive enzyme prior to secretion.

[0005] Although the functions of this enzyme have not been fully analysed, since the enzyme specifically cleaves the C-terminal of glutamic acid and aspartic acid, it is extensively used for determination of an amino acid sequence of proteins. In addition, since the present enzyme acts on a substrate even in the presence of urea (at a concentration of about 2M), it is used to liberate a desired peptide from its fusion protein, after solubilization, with urea, of a large amount of insoluble fusion protein intracellularly expressed according to a gene recombination technique.

[0006] The present inventors successfully used the above-mentioned method to efficiently produce human calcitonin by gene recombination techniques (Japanese Unexamined Patent Publication (Kokai) No. 5-328992, EP528686). In addition, the S. aureus V8 protease was used to excise human glucagon from a fusion protein expressed in the E. coli expression system (Kazumasa Yoshikawa et al. (1992), Journal of Protein Chemistry, 11, 517 - 525).

[0007] As can be seen from the above, the present enzyme has been extensively used for research and production of peptides by gene recombination. However, since the enzyme is purified from a culture medium of S. aureus V8, there are problems in that (1) the enzyme is contaminated with trace amounts of other proteins, (2) the S. aureus V8 is a prothogenic strain, and (3) the product is expensive.

[0008] It would be desirable to enable production of a large amount of a desired polypeptide such as S. aureus V8 protease. The production of highly purified desired polypeptides such as S. aureus V8 protease is very advantageous in scientific research and industry and a method for industrial production of large amounts of polypeptides has been urgently sought.

[0009] Accordingly, the present invention provides a process for the production of a desired polypeptide comprising the steps of:

(1) transforming host cells with an expression vector comprising a gene coding for a fusion protein comprising a desired polypeptide and a protective polypeptide;
(2) culturing the transformed host cells so as to express said gene to produce the fusion protein; and
(3) excising the desired polypeptide from the fusion protein with a protease intrinsic to the host cells.

[0010] According to a preferred embodiment of the present invention, there is provided a process for the production of a desired polypeptide, comprising the steps of:

(1) transforming Escherichia coli host cells with an expression vector comprising a gene coding for a fusion protein comprising at least one protective polypeptide, a desired polypeptide and a linker peptide, wherein the protective polypeptide is a polypeptide derived from E. coli β-galactosidase and/or a polypeptide derived from an amino glycoside 3'-phosphotransferase of transposone 903 origin, the desired polypeptide is a derivative of Staphylococcus aureus V8 protease, the linker peptide between said protective polypeptide and said desired polypeptide has a substrate site specifically recognized by a protease intrinsic to the host cells;
(2) expressing said gene in E. coli host cells to produce the derivative of the Staphylococcus aureus V8 protease as an inactive fusion protein;

(3) disrupting the cells so as to separate the fusion protein, and obtaining a fraction containing the E. coli ompT protease which is a protease intrinsic to the cells and the fusion protein;
(4) solubilizing the fusion protein with a denaturating agent; and
(5) decreasing a concentration of the denaturating agent to a level at which the E. coli ompT protease exhibits its activity to cleave the linker peptide with the protease so as to obtain the desired polypeptide from the fusion protein.

BRIEF EXPLANATION OF DRAWINGS

[0011] Fig. 1(a) and Fig. 1(b) represent (a) construction of a gene coding for Staphylococcus aureus (S. aureus) V8 protease, and (2) nucleotide sequence of PCR primes used to clone the gene of the present invention.
[0012] Fig. 2 represents processes for construction of pG97S4DhCT[G]R6 and pG97S4DhCT[G]R10.
[0013] Fig. 3 represents processes of construction of plasmid pV8RPT(+) and pV8RPT(-).
[0014] Fig. 4(a) and Fig. 4(b) represent amino acid sequences encoded in plasmids pV8RPT(+) and pV8RPT(-).
[0015] Fig. 5 represents a process for construction of plasmid pV8D.
[0016] Fig. 6 represents an amino acid sequence of a fusion protein encoded in the plasmid pV8D.
[0017] Fig. 7 represents a result of electrophoresis showing that the present fusion protein forms inclusion bodies and transfers to an insoluble fraction.
[0018] Fig. 8A and Fig. 8B represent a result of electrophoresis showing that the fusion protein V8D is cleaved by a host-derived protease ompT to liberate V8 protease.
[0019] Fig. 9 represents a result of electrophoresis showing refolding of V8 protease liberated by the ompT protease.
[0020] Fig. 10A and Fig. 10B are a chart comparing products formed from a fusion protein comprising a human calcitonin precursor by cleaving the fusion protein by (A) a recombinant V8 protease obtained by the present method and by (B) V8 protease obtained from S. aureus.
[0021] Fig. 11 represents nucleotide sequences of primers used for construction of DNAs coding for various fusion proteins (pV8H, pV8F, pV8A, pV8D2 and pV8Q).
[0022] Fig. 12 represents a process for construction of plasmids pV8H, pV8F, pV8A, pV8D2 and pV8Q.
[0023] Fig. 13 represents C-terminal amino acid sequences of V8 protease encoded in plasmids pV8D, pV8H, pV8F, pV8A, pV8D2 and pV8Q, as well as formation of inclusion bodies from expression products (fusion proteins) of the plasmids.
[0024] Fig. 14(a), Fig. 14(b) and Fig. 14(c) represent an amino acid sequence of a desired polypeptide produced by a process of the present invention.

DETAILED DESCRIPTION

[0025] To carry out the present invention, a gene coding for a desired polypeptide, for example a gene coding for S. aureus V8 protease, is isolated from, for example, S. aureus V8 or is synthesized, the gene thus obtained is introduced into safe host cells such as E. coli cells, and the desired polypeptide, such as an enzyme, is produced with a low cost.
[0026] Generally to produce a desired polypeptide or protein by genetic engineering, preferably the desired polypeptide or protein is formed as a fusion protein and the polypeptide or protein is intracellularly accumulated as insoluble inclusion bodies, so as to prevent the bad effects of the produced polypeptide or protein on the growth and survival of the host cell as well as expression of the desired polypeptide or protein.
[0027] However, in some cases according to a conventional procedure a fusion protein comprising a desired polypeptide or protein does not forms inclusion bodies. In such cases, according to the present invention both of the C-terminal and N-terminal of a desired polypeptide or protein are linked with protective peptides through linker peptides to form an insoluble fusion protein so as to form inclusion bodies.
[0028] According to an embodiment of the present invention, a fusion protein comprising a desired polypeptide or protein is intracellularly accumulated as inclusion bodies. Although a protease intrinsic to host cells does not act on the inclusion bodies, said protease accompanying the inclusion bodies acts on the fusion protein after the inclusion bodies are isolated and the fusion protein is dissolved, so as to cleave the fusion protein resulting in liberation of the desired polypeptide or protein. In this way, according to the present invention, a desired polypeptide or protein with high purity can be efficiently produced using genetic engineering.
[0029] As an example of the production of a desired polypeptide or protein according to the present invention, a process for a large-scale production of S. aureus V8 protease with high purity using the E. coli expression system is described in detail.
[0030] First, the present inventors considered that to produce a large amount of S. aureus V8 protease using an E. coli expression system, it is preferable to express the protease as an enzymatically inactive fusion protein, because it is considered that if the present protease is directly expressed in host cells the protease hydrolyses E. coli proteins resulting in termination of the growth of the host cells, and a large amount of S. aureus V8 protease cannot be obtained.

[0031] Accordingly, the present inventors planned to express the desired protein as an enzymatically inactive fusion protein, to excise an enzymatically active S. aureus V8 protease moiety from the fusion protein with another protease, to refold the excised V8 protease and to purify the active S. aureus V8 protease.

[0032] In addition, the present inventors considered the use of ompT protease considered to exist in the outer membrane of E. coli cell as the protease for cleaving the fusion protein, because it is considered that the addition of another enzyme increases the production cost.

[0033] The use of ompT protease intrinsic in E. coli cells has advantages in that (1) the production process is simple; (2) the addition of an additional enzyme into the reaction system is not necessary resulting in a low production cost; and (3) since the use of V8 protease separately produced by other host is not necessary, incorporation of proteins contaminated in the V8 protease preparation (for commercially available V8 protease, S. aureus-derived proteins) into the product can be prevented.

[0034] E. coli ompT protease exists in an E. coli outer membrane fraction and selectively cleaves a bond between two basic amino acids (Keijiro Sugimura and Tatsuro Nishihara (1988), J. Bacteriol. 170, 5625 - 5632). Sugimura et al. purified the ompT protease, subjected various peptides to cleavage with the ompT protease at 25°C for 30 minutes, and reported that the ompT protease cleaves a bond between two basic amino acids, i.e., a bond between arginine-arginine, lysine-arginine, lysine-lysine and lysine-arginine. However, they did not refer to whether or not the ompT protease is active in the presence of urea (urea concentration 2M or more). It is considered that since the ompT protein occurs in the outer membrane, the ompT protease coprecipitates with a precipitation fraction after producing S. aureus V8 protease as inclusion bodies in E. coli cells, disrupting the E. coli cells and centrifuging the disruptant to separate an insoluble fraction.

[0035] To the precipitated fraction thus obtained is added urea to solubilize the fusion protein comprising S. aureus V8 protein. The present inventors considered that if ompT protease activity is maintained in the above condition, the ompT protease could be used to excise S. aureus V8 protease from the fusion protein, and a large amount of S. aureus V8 protease could be produced by a simple process by refolding the excised protease to regenerate the active enzyme.

[0036] Next, the present invention is explained in detail. It was reported on the basis of a nucleotide sequence of S. aureus V8 protease gene that there are a signal sequence (pre sequence) necessary for secretion and a pro sequence, whose function is not clear, at the N-terminal of mature protein, and the above-mentioned repeat sequence at the C-terminal of the mature protein. Accordingly, a gene I coding for a mature protein from its N-terminal to C-terminal, and a gene II lacking the repeat sequence whose function is not known were prepared. Although it is not clear whether the repeat sequence whose function is not known is essential for enzyme activity, the present inventors considered that if this repeat sequence is not necessary, by deleting this repeat sequence resulting in lowering the molecular weight the number of molecules of the expressed protein per cell can increase resulting in an amount of the expressed protein.

[0037] Accordingly, chromosomal DNA preparation was isolated from Staphylococcus aureus V8 (ATCC 27733), and two V8 protease derivative genes I and II were prepared by PCR. To express these derivatives, plasmids pV8RPT (+) and pV8RPT(-) were constructed wherein an E. coli β-galactosidase derivative was used as a protective peptide in a fusion protein. In these plasmids a gene coding for E. coli β-galactosidase derivative and a gene coding for V8 protease derivative (I or II) are linked with a gene coding for a linker peptide containing arginine-arginine which is recognized and cleaved with ompT protease, under the regulation by a lactose promoter, to express a fusion protein.

[0038] It was considered that for the fusion protein thus designed it is possible to express a V8 protease derivative as insoluble fusion protein, to solubilize the fusion protein using urea, to cleave the fusion protein with ompT protease by decreasing urea concentration so as to separate the E. coli β-galactosidase derivative protein and the V8 protease derivative protein.

[0039] The plasmids as designed above were constructed, and induced to express the fusion protein in E. coli host cells with isopropyl-β-D-thio-galactopyranoside (IPTG). As a result the two fusion proteins did not become insoluble, and after disruption of the cells, enzyme activity was detected in the supernatants.

[0040] On the other hand, after induction with IPTG, the growth of the cells remarkably decreased. It was observed from an analysis by SDS polyacrylamide gel electrophoresis that intracellular proteins were degraded by enzymatic activity of expressed V8 protease derivative. Therefore, it was clarified that in a method wherein a fusion protein comprising the E. coli β-galactosidase derivative and V8 protease derivative is expressed, (1) an expressed enzyme has an enzymatic activity and inhibits the growth of host cells; and (2) since an amount of an expressed protein is very low said method is not suitable for the production of V8 protease derivative.

[0041] However, it was considered from the above-mentioned results that, (1) the pro-sequence of the N-terminal probably does not involve refolding of V8 protease; and (2) the repeat sequence is possibly not necessary for V8 protease activity because V8 protease derivative II not having the C-terminal repeat sequence was active.

[0042] Next, to produce an active V8 protease by expressing V8 protease as an insoluble fusion protein in an E. coli expression system, solubilizing the fusion protein with urea, liberating V8 protease from the fusion protein using a protease in the presence of urea and by refolding the liberated V8 protease, the present inventors started experiments on the basis of the following assumption.

(1) When the above-mentioned E. coli β-galactosidase was fused to the N-terminal of V8 protease, the resulting fusion protein was not insoluble. Therefore, the present inventors planned to add additional protective peptide to the above-mentioned fusion protein to form an insoluble fusion protein, and tried to use an aminoglycoside 3'-phosphotransferase protein desired from a kanamycine resistance gene of transposone 903 (Nucleic Acids Res. (1988) 16,358). Namely, the present inventors considered adding a part of aminoglycoside 3-phosphotransferase protein to the C-terminal of a fusion protein comprising a E. coli β-galactosidase derivative and V8 protease, through a linker peptide to promote insolubilization of a fusion protein so as to form inclusion bodies.

[0043] In addition, (2) the present inventors expected that if the above-mentioned R6 linkers are positioned at the N- and C-terminals of the V8 protease, the V8 protease can be liberated from the fusion protein by cleaving the fusion protein with the ompT protease which cleaves a bond between two basic amino acids.

[0044] Accordingly on the basis of the above-mentioned assumption, a novel expression plasmid, pV8D which expresses a fusion protein comprising a part of aminoglycoside 3'-phosphotransferase fused to an E. coli β-galactosidase derivative/V8 protease derivative fusion protein was constructed. Note that the V8 protein derivative encoded by this plasmid (designated V8D protein hereinafter) lacks C-terminal 8 amino acids in comparison to the above-mentioned V8 protease derivative II, and the N-terminal and C-terminal of this V8D have been fused to an E. coli β-galactosidase derivative and part of aminoglycoside 3'-phosphotransferase through the R6 linker peptides.

[0045] E. coli JM101 having pV8D was cultured, and induced with IPTG, and it was found from SDS PAGE that a fusion protein of about 60 kd thus expressed intracellularly formed insoluble inclusion bodies.

[0046] Next, the cultured cells were disrupted, and the disruptant was centrifuged to isolate inclusion bodies comprising a fusion protein, which were then dissolved by a denaturating agent such as urea, guanidine hydrochloride or a surfactant.

[0047] In Examples of the present invention, the inclusion bodies were dissolved with 8M urea, and the mixture was diluted to make the urea concentration 4M, and the whole was incubated at 37°C for 2 hours. It was confirmed that under this condition ompT protease intrinsic to E. coli cleaved the fusion protein and provided, in SDS PAGE analysis, 12 KDa, 26 KDa and 22 KDa bands corresponding to β-galactosidase derivative, V8D protein and a part of aminoglycoside 3'-phosphotransferase protein respectively.

[0048] On the other hand, when the same experiment was carried out using as host E. coli W3110M25 which is an ompT deficient mutant, the above-mentioned bands were not detected revealing that a protease which specifically cleaved the fusion protein was an ompT protease intrinsic to E. coli. In addition, to confirm that the fusion protein was specifically cleaved with ompT protease, a 26 KDa band corresponding to the V8D protein was extracted from the SDS-PAGE gel, and N-terminal amino acid sequence thereof was determined. As a result, it was. confirmed that a bond between arginine-arginine in the R6 linker peptide was cleaved.

[0049] Accordingly, it was found for the first time by the present inventors that both the fusion protein and the ompT protease were present in the inclusion bodies precipitated by centrifugation, and after solubilization of the inclusion bodies with 8M urea, the ompT protease was fully active in the presence of 4M urea and precisely cleaved the expected site of the amino acid sequence.

[0050] It is considered that the enzyme activity of the product is very low because the V8 protease derivative protein formed in the presence of urea was denatured. Therefore, the present inventors carried out refolding of the V8D proteins, if necessary, by lowering the concentration of the denaturing agent so as to determine whether an entimatically active V8D protein can be obtained. After the cleave reaction with the ompT protease, a sample was diluted 20-fold with 0.4M potassium phosphate buffer (pH 7.5), and allowed to stand overnight on ice. By this operation, about 20% of the V8D protein was refolded and recovered it enzymatic activity. After this operation, a sample was analysed by SDS-PAGE. As a result, after refolding, a major protein was the V8D protein.

[0051] The reason of this phenomenon is considered that although prior to refolding a β-galactosidase derivative, a part of aminoglycoside 3'-phosphotransferase protein and E. coli-derived proteins were present, after the refolding the V8D protease having protease activity hydrolized other accompanied proteins. This result is very advantageous for purification of the V8D protein after refolding.

[0052] Next, it was tested whether the V8D protease activated as described above has the same substrate specificity as that of native S. aureus V8 protease. A substrate (for example a fusion protein comprising a human calcitonine derivative) was reacted with a refolded V8D protease and a native enzyme at 30°C for an hour, and peptide fragments generated from a fusion protein by cleavage with the enzyme were analysed by a high performance liquid chromatography. As a result, elution patterns of the peptide fragments generated by both of the enzymes were same, revealing that the V8D protein prepared as descried above has the same substrate specificity as that of the native enzyme.

[0053] To intracellularly express the S. aureus V8 protease as insoluble inclusion bodies, said protease should not act on the fusion protein. On the other hand, after refolding, said protease should exhibit its enzymatic activity. Such apparently discrepant properties are requested to V8 protease derivative protein. For the V8D protease, a fusion protein comprising a part of native V8 protease starting from the N-terminal and ending at the 212nd amino acid was constructed, and the C-terminal of said V8 protease portion was extended and fused to a part of aminoglycoside 3'-phos-

photransferase protein to form a fusion protein, and the fusion protein was tested whether it forms inclusion bodies and whether it is reactivated by refolding.

[0054] Namely, the present inventors constructed expression plasmids pV8H, pV8F, pV8A, pV8D2 and pV8Q expressing fusion proteins of the V8 protease which C-terminal is extended by 2, 4, 6 and 8 amino acid residues respectively by PCR method and gene cloning. These plasmids were used to transform E. coli JM101 to construct transformats. The resulting transformats were cultured and induced with IPTG. As a result, the transformats having the plasmids pV8D, pV8H or pV8F formed inclusion bodies of fusion protein and after cleavage of the fusion protein by ompT protease, refolded enzymes were reactivated. On the other hand, other transformats did not form inclusion bodies, and after disrupting the cultured cells, the soluble fraction exhibited V8 protease activity. Accordingly, it was found that the E. coli V8 protease derivative protein should fuse at its 215th phenylalanine or an amino acid before the 215th amino acid with a protective polypeptide to form inclusion bodies of the fusion protein, followed by enzymatical cleavage of the fusion protein and refolding.

[0055] Although the present invention was explained by taking V8 protease as an example of desired polypeptide, the same principle and procedure can apply to other desired polypeptides or protein such as motilin, glucagon, adrenocorticotrophic hormone (ACTH), corticotropin-releasing hormone (CRH), secretin, growth hormone, insulin, growth hormone-releasing hormone (GRH), vasopressin, oxytocin, gastrin, glucagon-like peptide (GLP-1, GLP-2, 7-36 amide), cholecystokinin, vasoactive intestinal polypeptide (VIP), pituitary adenolate cyclase activating polypeptide (p.a.c.a.p.), gastrin releasing hormone, galanin, thyroid-stimulating hormone (TSH), luteinizing hormone-releasing hormone (LH-RH), calcitonin, parathyroid hormone (PTH, PTH(1-34), PTH(1-84), peptide histidine isoleucine (PHI), neuropeptide Y (nP.Y)), peptide YY (P.YY), pancreatic polypeptide (P.P.), somatostatin, TGF-α, TGF-β, nerve growth factor, fibroblast growth factor, relaxin, prolactin, atrial natriuretic peptide (ANP), B-type natriuretic peptide (BNP), C-type natriuretic peptide (CNP), angiotensin, or brain derived nutrient factor (BDNF), and further enzymes such as KEX 2 endoprotease to efficiently produce such desired polypeptides or proteins.

[0056] Accordingly, the present invention provides a process for production of a desired polypeptide characterized by transforming host cells with an expression vector containing a gene coding for a fusion protein comprising a protective polypeptide and a desired polypeptide, expressing the gene to produce the fusion protein and excising the desired polypeptide by a protease intrinsic to the host cells.

[0057] According to the present invention, the fusion protein can be represented by the formula (1) A-L-B or (2) A-L-B-L-C, wherein A and C represent protective polypeptides, B represents a desired polypeptide and L represents a linker peptide containing a substrate site recognized by a protease intrinsic to the host cells, and the fusion protein is cleaved in the linker peptide L so as to obtain the desired polypeptide from the fusion protein.

[0058] According to a preferred embodiment of the present invention, the desired polypeptide is a biologically or physiologically active polypeptide, preferably an enzyme, and more preferably proteolytic enzyme. In the most preferably embodiment of the present invention, the desired polypeptide is a protease, which is expressed in host cells as an inactive fusion protein, the host cells are disrupted to isolate the fusion protein which is then solubilized with a denaturating agent, and then the linker peptide region is cleaved with a protease intrinsic to the host cells to obtain the desired polypeptide from the fusion protein.

[0059] In another preferred embodiment, there is mentioned a process for production of a desired polypeptide wherein the desired polypeptide is a proteolytic enzyme, the desired polypeptide is expressed in host cells as an insoluble fusion protein comprising the desired polypeptide linked to a protective polypeptide through a linker peptide, the host cells are disrupted to isolate the fusion protein, the fusion protein is solubilized with a denaturating agent at a concentration at which a protease intrinsic to the host cells is not active, and the concentration of the denaturating agent is lowered to a level at which said intrinsic protease exhibits its enzymatic activity so that the intrinsic protease cleaves the linker peptide to obtain the desired polypeptide from the fusion protein. In this case, during the isolation process after the disruption of the cells, said intrinsic protease and the fusion protein preferably coexist in the same fraction.

[0060] The protective polypeptide may be any polypeptide which can be expressed as a part of a fusion protein comprising a desired polypeptide, and for example a polypeptide derived from E. coli β-galactosidase, a polypeptide derived from aminoglycaside 3'-phosphotransferase of transposon 903 origin etc. alone or in combination may be used.

[0061] Linker peptide is a peptide having a site specifically recognized by a protease intrinsic to the host cells which contain an expression vector for a desired polypeptide. A preferred embodiment of the linker peptide a polypeptide consisting of 2 to 50 amino acid resides and containing one or more pairs of two basic amino acid residues. A linker peptide may have the pairs of basic amino acid residues at both of the N- and C-terminal thereof.

[0062] The denaturating agent may be any substance which solubilizes a fusion protein and is for example urea, guanidine hydrochloride, surfactants etc. Urea is preferably used, and in this case a concentration of urea is preferably 1 to 8M. Concentration of urea after solubilization of the fusion protein may be any concentration at which a protease intrinsic to host cells exhibit its enzymatic activity.

[0063] In the present invention relating to a process for production of a desired polypeptide wherein the desired polypeptide is expressed as a fusion protein and the fusion protein is cleaved with a protease intrinsic to host cells,

the intrinsic protease and the desired polypeptide are not limixed anyway. Namely, the intrinsic protease may be any protease capable of processing a fusion protein after the fusion protein is expressed as an insoluble protein, and is for example E. coli ompT protease used in Example or the like. The desired polypeptide may be any polypeptide consisting of preferably 20 to 800 amino acid residues, and for example S. aureus protease and/or a derivative thereof as shown in the Examples hereinafter.

[0064] According to the present process for production of a desired polypeptide using gene recombination technique, especially a V8 protease derivative protein may be produced in an E. coli expression system. Namely, a preferred embodiment for production of a desired polypeptide comprises the steps of:

(1) transforming Escherichia coli host cells with an expression vector comprising a gene coding for a fusion protein comprising at least one protective polypeptide, a desired polypeptide and a linker peptide, wherein the protective polypeptide is a polypeptide derived from E. coli β-galactosidase and/or a polypeptide derived from an aminoglycoside 3'-phosphotransferase of transposone 903 origin, the desired polypeptide is a derivative of Staphylococcus aureus V8 protease, the linker peptide between said protective polypeptide and said desired polypeptide has a substrate site specifically recognized by a protease intrinsic to the host cells;
(2) expressing said gene in E. coli host cells to produce the derivative of the Staphylococcus aureus V8 protease as an inactive fusion protein;
(3) disrupting the cells so as to separate the fusion protein, and obtaining a fraction containing the E. coli ompT protease which is a protease intrinsic to the cells and the fusion protein;
(4) solubilizing the fusion protein with a denaturating agent; and
(5) decreasing a concentration of the denaturating agent to a level at which the E. coli ompT protease exhibits its activity to cleave the linker peptide with the protease so as to obtain the desired polypeptide from the fusion protein.

[0065] After refolding V8 protease derivative, this protein can be highly purified by conventional procedures for purification of protein, for example, gel filtration, ionic chromatography, hydrophobic chromatography. In addition, for the V8 protease derivative as shown in Examples, since after finishing the refolding reaction said derivative is the main proteinaceous component in the reaction mixture, the purification is very easy.

EXAMPLE

[0066] Next, the present invention is explained, in more detail, in following Examples.

Example 1 Isolation of S. aureus V8 protease gene

[0067] A V8 protease gene was prepared by PCR method on the basis of a reported nucleotide sequence. Three PCR primers shown in Fig. 1(b) were synthesized by a DNA sequencer (Applied Bio system). The primers I, II and III correspond to the regions of V8 protease gene and the primer I has at its 5'-terminal side a XhoI restriction enzyme site and the primers II and III have at their 5'-terminal side a SalI restriction enzyme site, as shown in Fig. 1. PCR was carried out using a chromosome prepared from Staphylococcus aureus V8 (ATCC 27733) by Jayaswal et al. method (J. Bacterial. 172: 5783 - 5788 (1990)) and the above-mentioned PCR primers. 2.5 units of Taq DNA polymerase was added to 50 μl of a reaction mixture containing 1.0 μM primers, 1 μg chromosomal DNA, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin, 200 μM dNTP (mixture of dATP, dCTP, dCTP and dTTP), and PCR of 90°C for 1 minute, 72°C for 2 minutes and 55°C for 2 minutes was carried out for 30 cycles.

[0068] As a result, a gene for a mature V8 protease containing a repeat sequence but not containing a prepro sequence (protease gene derivative I, 0.8 kb) was obtained by the primers I and II, and a gene for V8 protease containing neither prepro sequence nor repeat sequence (V8 protease gene derivative II, 0.7 kb) was obtained by the primers I and III. Next, these genes were subjected to electrophoresis, purified with SUPREP-2 (Takara Shuzo), and cleaved with restriction enzymes XhoI and SalI to obtain V8 protease derivative gene fragments I and II having cohesive ends XhoI and SalI.

Example 2 Construction of expression vectors pV8RPT(+) and pV8RPT(-) as well as expression of V8 protease derivative

[0069] A plasmid used in this Example, pG97S4DhCT[G]R6 is a plasmid which efficiently expresses a fusion protein comprising an E. coli β-galactosidase derivative and a human calcitonin precursor (hCT[G]), and can be constructed from plasmid pBR322 and plasmid pG97S4DhCT[G] (see, Japanese Unexamined Patent Publication No. 5-328992, EP528686, and Fig. 2). Escherichia coli W3110 containing the plasmid pG97S4DhCT[G] was designated Escherichia coli SBM323 and deposited with the National Institute of Bioscience and Human-Technology Agency of Industrial Sci-

ence and Technology, 1-3 Higashi, 1-chome, Tsukuba-shi, Ibaraki-ken, Japan, on August 8, 1991 as an international deposition under the Budapest Trealy as FERM BP-3503.

[0070] To express the V8 protease genes I and II obtained by PCR, the plasmid pG97S4DhCT[G]R6 was digested with XhoI and SalI, and a DNA fragment (3.1 kb) lacking a human calcitonin precursor gene was prepared by agarose gel electrophoresis. This DNA fragment was joined to the V8 protease gene fragment having XhoI and SalI cohesive ends as prepared above using T4 DNA ligase, and the ligation product was used to transform E. coli JM101 so as to construct a plasmid pV8RPT(+) containing the V8 protease gene derivative I, and a plasmid pV8RPT(-) containing the V8 protease gene derivative II (Fig. 3). As a host for the plasmid, E. coli JM101 (available, for example, from Takara Shuzo, Invitrogen Catalog No. c660-00 etc.) was used. Amino acid sequences of fusion proteins comprising a V8 protease derivative and a β-galactosidase derivative, expressed by the above-mentioned plasmids are shown in Fig. 4. E. coli JM101/pV8RPT(+) and E. coli JM101/pV8RPT(-) were separately cultured in 100 ml of LB medium (0.5% yeast extract, 1.0% Trypton, 0.5% NaCl) at 37°C until the absorbance OD660 reached 1.0, and the gene expression was induced by adding isopropylthiogalactopyranaside (IPTG) to the final concentration 2 mM. After the addition, culturing was further continued for 2 hours, and the culture was centrifuged to recover the microbial cells, which were then suspended in TE buffer (10 mM Tris-HCl(pH 8.0), 1 mM EDTA) at a concentration of OD660 = 5.

[0071] The cell suspension was treated with an ultrasonicator (cellruptor, Tosho Denski K.K.), the disruptant was centrifuged at 12,000 rpm for 5 minutes to eliminate an insoluble fraction, and the supernatant thus obtained was used as a crude enzyme preparation. The activity of V8 protease was measured using a synthetic substrate (Z-Phe-Leu-Glu-4-nitranilide; Boehringer Mamnnheim), 940 µl of 100 mM Tris-HCl (pH 8.0) buffer and 20 µl of 10 mM Z-Phe-Leu-Glu-4-nitranilide solution in DMSO were mixed, and to the mixture was added 40 µl of a crude enzyme solution. The mixture was incubated at room temperature for 5 minutes, and the absorbance, at 405 nm, of the reaction mixture was measured by a Hitachi spectrophotometer U-3200.

[0072] As a result, the crude enzyme solutions prepared from the cells of E. coli JM101/pV8RPT(+) and E. coli JM101/pV8RPT(-) provided an enzyme activity of 8 µg/ml, revealing that the enzyme in the form of a fusion protein with β-galactosidase and lacking a prepro sequence exhibits an enzyme activity. In addition, the product lacking C-terminal repeat sequence, encoded by pV8RPT(-) exhibited an enzyme activity, revealing that the repeat sequence is not essential for the enzyme activity.

[0073] Since the productivity of the V8 protease derivatives I and II by E. coli JM101/pV8RPT(-) and E. coli JM101/pV8RPT(+) was low and the bands thereof could not be detected by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), then purification thereof was difficult. Considering the facts that the growth of the cells was ceased by the addition of IPTG and that the content of high molecular weight protein in induced cells was lower than that in non-induced cells, it was considered that a cause of low productivity is the fatal toxicity of V8 protease intracellularly expressed.

Example 3 Construction of expression vector pV8D

[0074] As can be seen from the above, since V8 protease fused at its N-terminal with β-galactosidase derivative has still its enzyme activity, the V8 protease cannot be inactivated by fusion only at its N-terminal. Accordingly the present inventors attempted to inactivate the V8 protease by further fusing the C-terminal thereof with a part of aminoglycosede 3'-phosphotransferase. To fuse the C-terminal of the V8 protease, the EcoRV site positioned before the repeat sequence was used.

[0075] The plasmid pV8D coding for a V8 protease derivative was constructed according to the procedure as shown in Fig. 5. Namely, a BglII-SalI fragment (3.0 kb) and an EcoRV-BglII fragment (0.7 kb) were prepared from pV8RPT(-1), a NarI-SalI fragment (0.2 kb) was prepared from pG97S4DhCT[G]R10, and these three DNA fragments were joined to obtain pV8hCT[G]. Note that the pG97S4Dh[G]R10 can be constructed from plasmid pBR322 and plasmid pG97S4DhCT[G] according to the same procedure for the above-mentioned plasmid pG97S4DhCT[G]R6 (see Japanese Unexamined Patent publication No. 5-328992, EP528686, Fig. 2).

[0076] Next, a hCT[G] region (0.1 kb BstEII-SalI fragment) in the pV8hCT[G] obtained as described above was replaced with a 0.8 kb SmaI-SalI fragment containing an aminoglycoside 3'-phosphotransferase gene region derived from pUC4K (Vieira, J. and Messing, J., Gene 19,259 (1982), Pharmacia Ca. No. 27-4958-01) to construct pV8D. An amino acid sequence of a fusion protein (V8D fusion protein) comprising V8 protease derivative expressed by said plasmid pV8D is shown in Fig. 6.

[0077] This fusion protein comprises V8 protease linked at its N-terminal and C-terminal with a β-galactosidase derivative and a part of aminoglycoside 3'-phosphotransferase respectively, through R6 linkers. The R6 linker has the sequence RLYRRHHRWGRSGSPLRAHE (SEQ ID NO: 1) wherein the peptide bond between R-R will be cleaved with ompT protease of E. coli.

Example 4

[0078] According to a conventional procedure, E. coli JM101/pV8D transformed with pV8D was cultured in 100 ml of LB medium at 37°C until an absorbance OD660 reaches 0.6, and the production of fusion protein V8D was induced by adding IPTG to a final concentration 100 mM. After the addition, further culturing was continued for two hours, and the culture was centrifuged to recover the cultured cells. Contrary to the case of E. coli JM101/pV8RPT(+) and E. coli JM101/pV8RPT(-), in the case of the present strain, the growth of the cells was not terminated by the induction of expression of a fusion protein, and V8 protease activity was not detected in the cells.

[0079] In addition, the intracellular formation of inclusion bodies was microscopically observed. A result of 16% SDS-PAGE for cells as well as an insoluble fraction and a soluble fraction obtained by sonicating the cells, before and after the induction, is shown in Fig. 7. A large amount of V8D fusion protein of 60 kDa was contained in the induced cells, and in an insoluble fraction because inclusion bodies were formed. It is believed that the expression reached a level at which inclusion bodies were formed, because the V8 protease derivative fused at its C-terminal with a part of aminoglucoside 3'-phosphotransferase cannot take a native configuration and is inactive, and therefore it does not inhibit the growth of the cells. Note that in addition to the 60 kDa V8D fusion protein, a 27 kDa protein was detected in the insoluble fraction, and it was found that the 27 kDa protein is a fragment of the V8D fusion protein, which fragment comprised an amino acid sequence containing the 282nd methionine and following region.

Example 5 Processing of V8D fusion protein by ompT protease

[0080] Microbial cells obtained by culturing as shown in Example 4 were suspended in 10 ml of TE buffer, and disrupted by ultrasonic treatment. After that the inclusion bodies were recovered by centrifugation. The resulting inclusion bodies were re-suspended in 10 ml of deionized water, and the suspension was centrifuged to wash the inclusion bodies. The inclusion bodies were diluted with deionized water until the OD660 value reached 100, and 150 µl of the reaction mixture was taken. To the 150 µl sample were added 25 µl of 1M Tris-HCl (pH 8.0), 2.5 µl of 1M dithiothreitol (DTT) and 120 mg of urea to solubilize the inclusion bodies, and to the resulting solution was added deionized water to make total volume 500 µl. The resulting solution was heated at 37°C for 2 hours.

[0081] Fig. 8A shows a result of 16% SDS-PAGE before and after the heating. As can be seen from Fig. 8A, a sample after the heating provided bands corresponding to β-galactosidase derivative, V8 protease derivative and a part of aminoglucoside 3'-phosphotransferase having molecular weights of 12 kDa, 26 kDa and 22 kDa respectively. On the other hand, Fig. 8B shows a result obtained by using a protease deficient strain E. coli W3110M25 (Sugimura, K. (1987) Biochem. Biophys. Res. Commun. 153, 753 - 759), expressing the V8D fusion protein and treating the inclusion bodies thus obtained according to the same procedure as described above. In this case, the above-mentioned three bands were not detected in SDS-PAGE, and therefore it was determined that the processing of the present fusion protein was a specific cleavage by ompT protease (Sugimura, K. and Nishihara, T. (1988) J. Bacteriol., 170, 5625 - 5632).

[0082] In addition, excision of the 26 kDa band from the SDS-PAGE and determination of N-terminal amino acid sequence of the fragment revealed that the band between R-R in the R6 sequence (RLYRRHHRWGRSGSPLRAHE) (SEQ ID NO: 1) was cleaved, and it was confirmed that the cleavage of the fusion protein was specifically carried out by ompT protease. During the above-mentioned operation, the solubilization of the inclusion bodies was carried out in the presence of 8M urea, and the processing was carried out in the presence of 4M urea, and therefore it was shown for the first time that the ompT protease is resistant to such a high concentration of urea, and can specifically cleave the fusion protein solubilized from the inclusion bodies.

Example 6. Refolding of recombinant V8 protease (V8D)

[0083] A sample after processing was diluted 20-fold with 0.4M potassium phosphate buffer (pH 7.5), and allowed to stand overnight on ice. By this operation the recombinant V8 Protease (V8D) refolded, and exhibited an activity corresponding to 30 µg/ml as determined as described above. Ratio of refolding was about 20%. Fig. 9 shows a result of 16% SDS-PAGE before and after the refolding.

[0084] On refolding, the recombinant V8 protease (V8D) behaves as a strong protease on the β-galactosidase derivative, a protein derived from aminoglucoside 3'-phosphotransferase and other E. coli-derived proteins, which are therefore degraded and disappear. Accordingly, a sample after refolding contains the recombinant V8 protease (V8D) as a major protein, and therefore purification operation following the refolding may be extensively simplified.

[0085] The recombinant V8 protease (V8D) thus obtained lacks 56 amino acid residues at the C-terminal due to the construction of the gene, in comparison with the native V8 protease, but maintains its activity, revealing that this lacking region is not essential for the enzyme activity.

Example 7. Substrate specificity of recombinant V8 protease obtained by refolding from inclusion bodies

[0086] To compare substrate specificity of recombinant V8 protease obtained by refolding with that of native V8 protease, an experiment was carried out wherein both of the proteases act on a fusion protein of a calcitonin precursor (hCT[G]) as a substrate to liberate hCT[G]. The calcitonin fusion protein used in the experiment comprises a β-galactosidase derivative (108 amino acid residues) and hCT[G] linked through a linker having glutamic acid, and native V8 protease cleaves the peptide bond of the carboxy side of the glutamic acid to liberate hCT[G]. Note that as a plasmid coding for said fusion protein, pG97S4DhCT[G]R4 can be mentioned (see Japanese Unexamined Patent Publication No. 5-328992, EP528686).

[0087] An amount of recombinant V8 protease (V8D) corresponding to an activity of 1.2 μg of native V8 protease was added to 1 ml of a solution containing 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 5 mM DTT, 2M urea and 10 mg/ml human calcitonin fusion protein, the mixture was reacted at 30°C for one hour, and the reaction mixture was analyzed by high performance liquid chromatography wherein elution was carried out using 0.1% trifluoroacetic acid (TFA) and 0.1% TFA/50% acetonitrile. Note that the recombinant V8 protease (V8D) was used after refolding without further purification, and a commercially available native V8 protease was used as a control.

[0088] Fig. 10 shows an elution pattern of the high performance liquid chromatography.

[0089] Cleavage patterns of the recombinant V8 protease and the native V8 protease on the human calcitonin fusion protein were same, confirming that substrate specificity of both of the proteases is same.

Example 8. Study of fusion site at C-terminal side of V8 protease

[0090] To construct fusion proteins of V8 protease whose C-terminal is extended by 2, 3, 4, 6 and 8 amino acid residues respectively, in comparison with the V8D fusion protein, PCR primers shown in Fig. 11 were synthesized. A plasmid pV8F coding for a fusion protein extended by 3 amino acid residues (V8F fusion protein) was constructed as follows.

[0091] Namely, V8 protease gene was amplified using the primer b and the primer I shown in Fig. 1(b), and using as a template 0.1 μg of pV8RPT(-) constructed in Example 1, the amplified DNA fragment was cleaved with EcoRI and SacI to prepare a 0.1 kb gene fragment. On the other hand, a R6 linker sequence and aminoglucoside 3'-phosphotransferase gene region was amplified using the primer g and primer h, and 0.1 μg of pV8D as a template DNA, and the amplified DNA fragment was cleaved with EcoT22I and SacI to prepare a 0.3 kb gene fragment. Note that the PCR was carried out under the same condition as described in Example 1. The 0.1 kb and 0.3 kb gene fragments obtained as described above and the EcoRI-EcoT22I fragment (4.2 Kb) of pV8D were joined to construct pV8F (see, Figs. 12 and 14).

[0092] Plasmid pV8H, pV8A, pV8D2 and pV8Q were constructed using primer g and primers a, c, d and e respectively according to the same procedure as described above (except that NdeI was used in place of SacI for construction of pV8H). Combinations of primers and template DNAs used for construction of the above plasmids are as follow.

pV8H: primer a, primer I and pV8RPT(-) as well as PCR product obtained by a combination of primer f, primer h and pV8D;
pV8A: primer c, primer I and pV8RPT(-) as well as PCR product obtained by a combination of primer g, primer h and pV8D;
pV8D2: primer d, primer I and pV8RPT(-) as well as PCR product obtained by a combination of primer g, primer h and pV8D; and
pV8Q: primer e, primer I and pV8RPT(-) as well as PCR product obtained by combination of primer g, primer h and pV8D.

[0093] These plasmids produce fusion proteins comprising V8 protease region whose C-terminal is extended by 2, 4, 6 and 8 amino acid residues respectively in comparison with the V8D fusion protein shown in Example 4 (see, Figs. 13 and 14).

[0094] These plasmids were used to transform E. coli JM101, and expression of each fusion protein was tested according to the same procedure as described in Example 4. The result is shown in Fig. 13. Inclusion bodies were formed from pV8H, pV8F and pV8D, and the inclusion bodies thus obtained were converted to refolded active V8 proteases according to the same procedure as described in Example 5. On the other hand, pV8A, pV8D2 and pV8Q did not form inclusion bodies, and V8 protease activity was detected in a soluble fraction. In the case of these plasmids, it is considered that the inclusion bodies are not formed because the expressed fusion proteins have protease activity and inhibit the growth of host cells. Namely, it was found that to express V8 protease as an inactive fusion protein, it is important to fuse V8 protease at its 215th phenylalanine or an amino acid before (nearer to the N-terminal) said phenylalanine with a protective polypeptide, and if the V8 protease fuses at an amino acid after (nearer to the C-

terminal) the 215th phenylalanine, since a fusion protein whose V8 protease moiety forms a native configuration exhibiting protease activity is produced, the growth is repressed resulting in low expression.

[0095] According to the present invention, a large amount of a desired polypeptide can be produced with a low cost. Especially, according to the present invention, a large amount of S. aureus V8 protease can be efficiently produced with low cost using safe host such as E. coli according to gene recombination procedures.

SEQUENCE LISTING

[0096]

SEQ ID NO: 1

SEQUENCE LENGTH: 20
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Peptide

SEQUENCE

```
Arg Leu Tyr Arg Arg His His Arg Trp Gly Arg Ser Gly Ser Pro Leu
 5                  10                  15

Arg Ala His Glu
            20
```

SEQ ID NO: 2

SEQUENCE LENGTH: 31
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

```
ACCGCTCGAG GTTATATTAC CAAATAACGA T                                  31
```

SEQ ID NO: 3

SEQUENCE LENGTH: 30
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

```
CTTAATGTCG ACTTAAGCTG CATCTGGATT                                    30
```

SEQ ID NO: 4

SEQUENCE LENGTH: 31
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

TCGCGTCGAC TTATTGGTCA TCGTTGGCAA A 31

SEQ ID NO: 5

SEQUENCE LENGTH: 344
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

```
Met Thr Met Ile Thr Asp Ser Leu Ala Val Val Leu Gln Arg Arg Asp
 1               5                  10                  15
Trp Glu Asn Pro Gly Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
            20                  25                  30
Pro Phe Ala Ser Trp Arg Asn Ser Asp Asp Ala Arg Thr Asp Arg Pro
        35                  40                  45
Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe
    50                  55                  60
Pro Ala Pro Glu Ala Val Pro Asp Ser Leu Leu Asp Ser Asp Leu Pro
65                  70                  75                  80
Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp Gln Met His Gly Tyr
                85                  90                  95
Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His Arg Trp Gly Arg Ser
            100                 105                 110
```

```
Gly Ser Pro Leu Arg Ala His Glu Gln Phe Leu Glu Val Ile Leu Pro
        115                 120                 125
Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn Gly His Tyr Ala
    130                 135                 140
Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly Thr Phe Ile Ala
145                 150                 155                 160
Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr Asn Lys His Val
                165                 170                 175
Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys Ala Phe Pro Ser
            180                 185                 190
Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe Thr Ala Glu Asn
        195                 200                 205
Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile Val Lys Phe Ser
    210                 215                 220
Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val Lys Pro Ala Thr
225                 230                 235                 240
Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn Ile Thr Val Thr
                245                 250                 255
Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp Glu Ser Lys Gly
            260                 265                 270
Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr Asp Leu Ser Thr
        275                 280                 285
Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu Lys Asn Glu Val
    290                 295                 300
Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe Asn Gly Ala Val
305                 310                 315                 320
Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln Asn Ile Glu Asp
                325                 330                 335
Ile His Phe Ala Asn Asp Asp Gln
            340
```

SEQ ID NO: 6

SEQUENCE LENGTH: 392
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear

MOLECULE TYPE: Polypeptide

SEQUENCE

```
Met Thr Met Ile Thr Asp Ser Leu Ala Val Val Leu Gln Arg Arg Asp
 1               5                  10                  15
Trp Glu Asn Pro Gly Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
            20                  25                  30
Pro Phe Ala Ser Trp Arg Asn Ser Asp Asp Ala Arg Thr Asp Arg Pro
        35                  40                  45
Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe
    50                  55                  60
Pro Ala Pro Glu Ala Val Pro Asp Ser Leu Leu Asp Ser Asp Leu Pro
65                  70                  75                  80
Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp Gln Met His Gly Tyr
                85                  90                  95
Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His Arg Trp Gly Arg Ser
            100                 105                 110
Gly Ser Pro Leu Arg Ala His Glu Gln Phe Leu Glu Val Ile Leu Pro
        115                 120                 125
Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn Gly His Tyr Ala
    130                 135                 140
Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly Thr Phe Ile Ala
145                 150                 155                 160
Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr Asn Lys His Val
                165                 170                 175
Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys Ala Phe Pro Ser
            180                 185                 190
```

```
Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe Thr Ala Glu Asn
        195                 200                 205
Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile Val Lys Phe Ser
    210                 215                 220
Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val Lys Pro Ala Thr
225                 230                 235                 240
Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn Ile Thr Val Thr
                245                 250                 255
Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp Glu Ser Lys Gly
            260                 265                 270
Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr Asp Leu Ser Thr
        275                 280                 285
Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu Lys Asn Glu Val
    290                 295                 300
Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe Asn Gly Ala Val
305                 310                 315                 320
Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln Asn Ile Glu Asp
                325                 330                 335
Ile His Phe Ala Asn Asp Asp Gln Pro Asn Asn Pro Asp Asn Pro Asp
            340                 345                 350
Asn Pro Asn Asn Pro Asp Asn Pro Asn Asn Pro Asp Glu Pro Asn Asn
        355                 360                 365
Pro Asp Asn Pro Asn Asn Pro Asp Asn Pro Asp Asn Gly Asp Asn Asn
    370                 375                 380
Asn Ser Asp Asn Pro Asp Ala Ala
385                 390
```

SEQ ID NO: 7

SEQUENCE LENGTH: 532
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

```
Met Thr Met Ile Thr Asp Ser Leu Ala Val Val Leu Gln Arg Arg Asp
 1               5                  10                  15
Trp Glu Asn Pro Gly Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
            20                  25                  30
Pro Phe Ala Ser Trp Arg Asn Ser Asp Asp Ala Arg Thr Asp Arg Pro
        35                  40                  45
Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe
    50                  55                  60
Pro Ala Pro Glu Ala Val Pro Asp Ser Leu Leu Asp Ser Asp Leu Pro
65                  70                  75                  80
Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp Gln Met His Gly Tyr
                85                  90                  95
Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His Arg Trp Gly Arg Ser
            100                 105                 110
Gly Ser Pro Leu Arg Ala His Glu Gln Phe Leu Glu Val Ile Leu Pro
        115                 120                 125
Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn Gly His Tyr Ala
    130                 135                 140
Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly Thr Phe Ile Ala
145                 150                 155                 160
Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr Asn Lys His Val
                165                 170                 175
Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys Ala Phe Pro Ser
            180                 185                 190
Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe Thr Ala Glu Asn
        195                 200                 205
Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile Val Lys Phe Ser
    210                 215                 220
```

```
Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val Lys Pro Ala Thr
225                 230                 235                 240
Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn Ile Thr Val Thr
                245                 250                 255
Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp Glu Ser Lys Gly
            260                 265                 270
Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr Asp Leu Ser Thr
        275                 280                 285
Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu Lys Asn Glu Val
    290                 295                 300
Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe Asn Gly Ala Val
305                 310                 315                 320
Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln Asn Ile Glu Asp
                325                 330                 335
Arg Leu Tyr Arg Arg His His Arg Trp Gly Arg Ser Gly Ser Pro Leu
            340                 345                 350
Arg Ala His Glu Gln Phe Leu Glu Cys Gly Asn Gly Lys Thr Ala Phe
        355                 360                 365
Gln Val Leu Glu Glu Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala
    370                 375                 380
Leu Ala Val Phe Leu Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys
385                 390                 395                 400
Pro Phe Asn Ser Asp Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg
                405                 410                 415
Met Asn Asn Gly Leu Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn
            420                 425                 430
Gly Trp Pro Val Glu Gln Val Trp Lys Glu Met His Lys Leu Leu Pro
        435                 440                 445
Phe Ser Pro Asp Ser Val Val Thr His Gly Asp Phe Ser Leu Asp Asn
    450                 455                 460
```

```
Leu Ile Phe Asp Glu Gly Lys Leu Ile Gly Gly Ile Asp Val Gly Arg
465                 470                 475                 480

Val Gly Ile Ala Asp Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys
                485                 490                 495

Leu Gly Glu Phe Ser Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr
            500                 505                 510

Gly Ile Asp Asn Pro Asp Met Asn Lys Leu Gln Phe His Leu Met Leu
        515                 520                 525

Asp Glu Phe Phe
    530
```

SEQ ID NO: 8

SEQUENCE LENGTH: 30
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

```
ATCGTTGGCC ATATGGATAT CTTCAATATT                                    30
```

SEQ ID NO: 9

SEQUENCE LENGTH: 33
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

```
GACTTATTGG TCATCGAGCT CAAAATGGAT ATC                                33
```

SEQ ID NO: 10

SEQUENCE LENGTH: 30
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

GACTTATTGG TCGAGCTCGG CAAAATGGAT 30

SEQ ID NO: 11

SEQUENCE LENGTH: 33
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

ATCTGGGTTG AGCTCATCGT TGGCAAAATG GAT 33

SEQ ID NO: 12

SEQUENCE LENGTH: 33
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

ATCTGGTTGG AGCTCTTGGT CATCGTTGGC AAA 33

SEQ ID NO: 13

SEQUENCE LENGTH: 34
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

ACAAAATCAT ATGGAACGCC TATATCGCCG ACAT 34

SEQ ID NO: 14

SEQUENCE LENGTH: 33
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

AATATTGAAG AGCTCCGCCT ATATCGCCGA CAT 33

SEQ ID NO: 15

SEQUENCE LENGTH: 27
SEQUENCE TYPE: Nucleic acid
STRANDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Synthetic DNA

SEQUENCE

GAATGGCAAA AGCTTATGCA TTTCTTT 27

SEQ ID NO: 16

SEQUENCE LENGTH: 12
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Arg Leu Tyr Arg Arg His His Arg
5 10

SEQ ID NO: 17

SEQUENCE LENGTH: 16
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Ile His Met Glu Arg Leu Tyr Arg Arg His His Arg
5 10 15

SEQ ID NO: 18

SEQUENCE LENGTH: 17
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Ile His Phe Glu Leu Arg Leu Tyr Arg Arg His His
5 10 15

Arg

SEQ ID NO: 19

SEQUENCE LENGTH: 18
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Ile His Phe Ala Glu Leu Arg Leu Tyr Arg Arg His
5 10 15

His Arg

SEQ ID NO: 20

SEQUENCE LENGTH: 19
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Ile His Phe Ala Asn Asp Glu Leu Arg Leu Tyr Arg
5 10 15

Arg His His Arg
20

SEQ ID NO: 21

SEQUENCE LENGTH: 22
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

Asn Ile Glu Asp Ile His Phe Ala Asn Asp Asp Gln Glu Leu Arg Leu
5 10 15

Tyr Arg Arg His His Arg
20

SEQ ID NO: 22

SEQUENCE LENGTH: 213
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

```
Val Ile Leu Pro Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn
 1               5                  10                  15

Gly His Tyr Ala Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly
            20                  25                  30

Thr Phe Ile Ala Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr
        35                  40                  45

Asn Lys His Val Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys
    50                  55                  60

Ala Phe Pro Ser Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe
65                  70                  75                  80

Thr Ala Glu Asn Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile
                85                  90                  95

Val Lys Phe Ser Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val
            100                 105                 110

Lys Pro Ala Thr Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn
        115                 120                 125

Ile Thr Val Thr Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp
    130                 135                 140

Glu Ser Lys Gly Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr
145                 150                 155                 160

Asp Leu Ser Thr Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu
                165                 170                 175

Lys Asn Glu Val Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe
            180                 185                 190

Asn Gly Ala Val Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln
        195                 200                 205

Asn Ile Glu Asp Ile
    210
```

SEQ ID NO: 23

SEQUENCE LENGTH: 214
SEQUENCE TYPE: Amino acid
TOPOLOGY: Linear

MOLECULE TYPE: Polypeptide

SEQUENCE

```
Val Ile Leu Pro Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn
 1               5                  10                  15
Gly His Tyr Ala Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly
            20                  25                  30
Thr Phe Ile Ala Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr
        35                  40                  45
Asn Lys His Val Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys
    50                  55                  60
Ala Phe Pro Ser Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe
65                  70                  75                  80
Thr Ala Glu Asn Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile
                85                  90                  95
Val Lys Phe Ser Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val
            100                 105                 110
Lys Pro Ala Thr Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn
        115                 120                 125
Ile Thr Val Thr Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp
    130                 135                 140
Glu Ser Lys Gly Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr
145                 150                 155                 160
Asp Leu Ser Thr Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu
                165                 170                 175
Lys Asn Glu Val Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe
            180                 185                 190
Asn Gly Ala Val Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln
        195                 200                 205
Asn Ile Glu Asp Ile His
    210
```

SEQ ID NO: 24

SEQUENCE LENGTH: 215
SEQUENCE TYPE: Amino acid

TOPOLOGY: Linear
MOLECULE TYPE: Polypeptide

SEQUENCE

```
Val Ile Leu Pro Asn Asn Asp Arg His Gln Ile Thr Asp Thr Thr Asn
 1               5                  10                  15
Gly His Tyr Ala Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly
            20                  25                  30
Thr Phe Ile Ala Ser Gly Val Val Val Gly Lys Asp Thr Leu Leu Thr
        35                  40                  45
Asn Lys His Val Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys
    50                  55                  60
Ala Phe Pro Ser Ala Ile Asn Gln Asp Asn Tyr Pro Asn Gly Gly Phe
65                  70                  75                  80
Thr Ala Glu Asn Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile
                85                  90                  95
Val Lys Phe Ser Pro Asn Glu Gln Asn Lys His Ile Gly Glu Val Val
            100                 105                 110
Lys Pro Ala Thr Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn
        115                 120                 125
Ile Thr Val Thr Gly Tyr Pro Gly Asp Lys Pro Val Ala Thr Met Trp
    130                 135                 140
Glu Ser Lys Gly Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr
145                 150                 155                 160
Asp Leu Ser Thr Thr Gly Gly Asn Ser Gly Ser Pro Val Phe Asn Glu
                165                 170                 175
Lys Asn Glu Val Ile Gly Ile His Trp Gly Gly Val Pro Asn Glu Phe
            180                 185                 190
```

```
Asn Gly Ala Val Phe Ile Asn Glu Asn Val Arg Asn Phe Leu Lys Gln
        195                 200                 205

Asn Ile Glu Asp Ile His Phe
    210                 215
```

Claims

1. A process for the production of a desired polypeptide comprising the steps of:

   transforming host cells with an expression vector comprising a gene coding for a fusion protein comprising the desired polypeptide and a protective polypeptide;
   culturing the transformed cells so as to express said gene to produce a fusion protein which is insoluble and deposited intracellularly as inclusion bodies;
   disrupting said host cells to separate said inclusion bodies, and
   using a protease to excise the desired polypeptide from the fusion protein and produce the desired polypeptide;

   characterised in that the protease used to excise the desired polypeptide from the fusion protein is endogenous protease from the disrupted host cells themselves, separated together with the inclusion bodies.

2. A process according to claim 1 comprising solubilising the fusion protein with a denaturation agent and cleaving the linker peptide by decreasing the concentration of the denaturation agent to a lower concentration at which the protease endogenous to the host cells exhibits enzymatic activity, so as to obtain the desired polypeptide from the fusion protein.

3. A process according to claim 2 wherein the desired peptide is refolded into an active conformation after cleavage of the fusion protein by the protease by decreasing the concentration of denaturing agent.

4. A process according to any one of the preceding claims wherein the fusion protein is represented by the formula (1) A-L-B, or (2) A-L-B-L-C, wherein A and C are protective polypeptides, B is a desired polypeptide and L is a linker peptide containing a substrate site specifically recognized by a protease intrinsic to the host cells, and the fusion protein is cleaved in the linker peptide L region so as to obtain the desired polypeptide B.

5. A process according to any one of the preceding claims wherein the host cells are E. coli.

6. A process according to any one of the preceding claims wherein said desired peptide is physiologically active and selected from the group consisting of motilin, glucagon, adrenocorticotrophic hormone (ACTH), corticotropin-releasing hormone (CRH), secretin, growth hormone, insulin, growth hormone-releasing hormone (GRH), vasopressin, oxytocin, gastrin, glucagon-like peptide (CLP-1, GLP-2, 7-36 amide), cholecystokinin, vasoactive intestinal polypeptide (VIP), pituitary adenolate cyclase activating polypeptide (p.a.c.a.p.), gastrin releasing hormone, galanin, thyroid-stimulating hormone (TSH), luteinizing hormone-releasing hormone (LH-RH), calcitonin, parathyroid hormone (PTH, PTH(1-34), PTH(1-84), peptide histidine isoleucine (PHI), neuropeptide Y (nP.Y)), peptide YY (P.YY), pancreatic polypeptide (P.P.), somatostatin, TGF-α, TGF-β, nerve growth factor, fibroblast growth factor, relaxin, prolactin, natriuretic peptide, angiotensin, and brain derived nutrient factor.

7. A process according to claim 5 wherein the natriuretic peptide is selected from the group consisting of ANP, BNP and CNP.

8. A process according to any one of claims 1 to 5 wherein the desired polypeptide is an enzyme.

9. A process according to claim 8 wherein the desired polypeptide is a proteolytic enzyme.

10. A process according to claim 9 wherein the proteolytic enzyme is selected from the group consisting of KEX2 endopeptidase, Staphylococcus aureus V8 protease and a derivative of Staphylococcus aureus V8 protease.

11. A process according to any one of the preceding claims wherein said desired polypeptide has 20 to 800 amino acid residues.

12. A process according to claim 10 wherein the S. aureus V8 protease derivative has the amino acid sequence indicated by underlining in Fig 4, or Fig 6, or an amino acid sequence as shown in Fig 14.

13. A process according to any one of claims 2 to 12 wherein the denaturation agent is selected from the group consisting of urea, guanidine hydrochloride and surfactants.

14. A process according to claim 13 wherein denaturation agent is 1 to 8 M urea.

15. A process according to claim 14 wherein the concentration of urea during the cleavage of the fusion protein is from 1 M to 6 M.

16. A process according to any one of the preceding claims wherein protective polypeptide is derived from E. coli β-gal and/or aminoglycoside 3'-phosphotransferase of transposon 903.

17. A process according to any one of claims 4 to 16 wherein protective polypeptide A is derived from E. coli β-gal and protective polypeptide C is derived from aminoglycoside 3'-phosphotransferase of transposon 903.

18. A process according to any one of the preceding claims wherein the protease is the E. coli ompT protease.

19. A process according to any one of the preceding claims wherein the protease is ompT and the fusion protein is cleaved with the protease in a solution containing about 4M urea.

20. A process according to any one of claims 4 to 19 wherein the linker peptide has 2 to 50 amino acids and one or two basic pairs of amino acids.

21. A process according to any one of the claims 4 to 20 wherein said linker peptide has basic pairs of amino acids at the C-terminal and N-terminal.

22. A process according to any one of claims 4 to 21 wherein the linker peptide has the amino acid sequence

RLYRRHHRWGRSGSPLRAHE(seq id no 1).

23. A process for production of a desired polypeptide, comprising the steps of:

(1) transforming Escherichia coli host cells with an expression vector comprising a gene coding for a fusion protein comprising at least one protective polypeptide, a desired polypeptide and a linker peptide, wherein the protective polypeptide is a polypeptide derived from E. coli β-galactosidase and/or a polypeptide derived from an aminoglycoside 3' - phosphotransferase of transposon 903 origin, the desired polypeptide is a derivative of Staphylococcus aureus V8 protease, the linker peptide between said protective polypeptide and said desired polypeptide has a substrate site specifically recognized by a protease intrinsic to the host cells;
(2) expressing said gene in E. coli host cells to produce the derivative of the Staphylococcus aureus V8 protease as an inactive fusion protein;
(3) disrupting the cells so as to separate the fusion protein, and obtaining a fraction containing the E. coli ompT protease which is an endogenous protease of the host cells and the fusion protein;
(4) solubilizing the fusion protein with a denaturation agent; and
(5) decreasing a concentration of the denaturation agent to a level at which the E. coli ompT protease exhibits its activity to cleave the linker peptide with the protease so as to obtain the desired polypeptide from the fusion protein.

Patentansprüche

1. Verfahren zur Herstellung eines gewünschten Polypeptids, welches die Schritte umfaßt:

Transformieren von Wirtszellen mit einem Expressionsvektor, der ein Gen umfaßt, das ein Fusionsprotein codiert, das das gewünschte Polypeptid und ein schützendes Polypeptid umfaßt;
Züchten der transformierten Zellen, um das Gen auszuprägen, wodurch ein Fusionsprotein erzeugt wird, das unlöslich ist und intrazellulär als Einschlußkörper abgelagert wird;
Aufbrechen der Wirtszellen, um die Einschlußkörper abzutrennen, und
Verwenden einer Protease, um das gewünschte Polypeptid aus dem Fusionsprotein auszuschneiden und das gewünschte Polypeptid zu erzeugen;

dadurch gekennzeichnet, daß die zum Ausschneiden des gewünschten Polypeptids aus dem Fusionsprotein verwendete Protease endogene Protease aus den aufgebrochenen Wirtszellen selbst ist, die zusammen mit den Einschlußkörpern abgetrennt wurde.

2. Verfahren nach Anspruch 1, das das Lösen des Fusionsproteins mit einem Denaturierungsmittel und das Abspalten eines Linker-Peptids umfaßt, indem die Konzentration des Denaturierungsmittels auf eine geringere Konzentration verringert wird, bei der die für die Wirtszellen endogene Protease eine enzymatische Aktivität zeigt, so daß das gewünschte Polypeptid aus dem Fusionsprotein erhalten wird.

3. Verfahren nach Anspruch 2, wobei bei dem gewünschten Peptid eine Rückfaltung in die aktive Konformation vorgenommen wird, nachdem das Fusionsprotein durch die Verringerung der Konzentration des Denaturierungsmittels durch die Protease gespalten wurde.

4. Verfahren nach einem der vorstehenden Ansprüche, wobei das Fusionsprotein mit der Formel (1) A-L-B oder (2) A-L-B-L-C angegeben wird, worin A und C schützende Polypeptide sind, B das gewünschte Polypeptid ist und L ein Linker-Peptid ist, das eine Substratstelle enthält, die von einer für die Wirtszellen intrinsischen Protease spezifisch erkannt wird, und das Fusionsprotein in der L-Region des Linker-Peptids abgespalten wird, so daß das gewünschte Polypeptid B erhalten wird.

5. Verfahren nach einem der vorstehenden Ansprüche, wobei die Wirtszellen E. coli sind.

6. Verfahren nach einem der vorstehenden Ansprüche, wobei das gewünschte Peptid physiologisch aktiv ist und aus der Gruppe ausgewählt ist, die aus Motilin, Glucagon, einem adrenocorticotropen Hormon (ACTH), einem Corticotropin freisetzenden Hormon (CRH), Secretin, einem Wachstumshormon, Insulin, einem ein Wachstumshormon freisetzenden Hormon (GRH), Vasopressin, Oxytocin, Gastrin, einem Glucagon ähnlichen Peptid (GLP-1, GLP-2, 7-36 Amid), Cholecystokinin, einem vasoaktiven, intestinalen Polypeptid (VIP), einem hypophysäre Adenolat-Cyclase aktivierenden Polypeptid (p.a.c.a.p.), einem Gastrin freisetzenden Hormon, Galanin, einem Thyroid stimulierenden Hormon (TSH), einem ein luteinisierendes Hormon freisetzenden Hormon (LH-RH), Calcitonin, einem parathyroiden Hormon (PTH, PTH(1-34), PTH(1-84), Peptid-Histidin-Isoleucin (PHI), Neuropeptid Y (nP.Y)), Peptid YY (P.YY), einem pankreatischen Polypeptid (P.P.), Somatostatin, TGF-α, TGF-β, einem Nervenwachstumsfaktor, einem Fibroblasten-Wachstumsfaktor, Relaxin, Prolactin, einem natriuretischen Peptid, Angiotensin und einem vom Hirn stammenden Nährfaktor besteht.

7. Verfahren nach Anspruch 5, wobei das natriuretische Peptid aus der Gruppe ausgewählt ist, die aus ANP, BNP und CNP besteht.

8. Verfahren nach einem der Ansprüche 1 bis 5, wobei das gewünschte Polypeptid ein Enzym ist.

9. Verfahren nach Anspruch 8, wobei das gewünschte Polypeptid ein proteolytisches Enzym ist.

10. Verfahren nach Anspruch 9, wobei das proteolytische Enzym aus der Gruppe ausgewählt ist, die aus KEX2-Endopeptidase, Staphylococcus aureus V8-Protease und einem Derivat von Staphylococcus aureus V8-Protease besteht.

11. Verfahren nach einem der vorstehenden Ansprüche, wobei das gewünschte Polypeptid 20 bis 800 Aminosäurereste hat.

12. Verfahren nach Anspruch 10, wobei das Derivat von S. aureus V8-Protease die Aminosäuresequenz, die in Figur 4 oder Figur 6 durch Unterstreichen gekennzeichnet ist, oder eine Aminosäuresequenz hat, wie sie in Figur 14 gezeigt ist.

13. Verfahren nach einem der Ansprüche 2 bis 12, wobei das Denaturierungsmittel aus der Gruppe ausgewählt ist, die aus Harnstoff, Guanidinhydrochlorid und oberflächenaktiven Mitteln besteht.

14. Verfahren nach Anspruch 13, wobei das Denaturierungsmittel 1 bis 8 m Harnstoff ist.

15. Verfahren nach Anspruch 14, wobei die Konzentration des Harnstoffs bei der Spaltung des Fusionsproteins 1 m bis 6 m beträgt.

16. Verfahren nach einem der vorstehenden Ansprüche, wobei das schützende Polypeptid von E. coli β-Gal und/oder Aminoglycosid-3'-Phosphotransferase von Transposon 903 stammt.

17. Verfahren nach einem der Ansprüche 4 bis 16, wobei das schützende Polypeptid A von E. coli β-Gal stammt und das schützende Polypeptid C von Aminoglycosid-3'-Phosphotransferase von Transposon 903 stammt.

18. Verfahren nach einem der vorstehenden Ansprüche, wobei die Protease E. coli ompT-Protease ist.

19. Verfahren nach einem der vorstehenden Ansprüche, wobei die Protease ompT ist und das Fusionsprotein mit der Protease in einer Lösung abgespalten wird, die etwa 4 m Harnstoff enthält.

20. Verfahren nach einem der Ansprüche 4 bis 19, wobei das Linker-Peptid 2 bis 50 Aminosäuren und ein oder zwei basische Aminosäurenpaare hat.

21. Verfahren nach einem der Ansprüche 4 bis 20, wobei das Linker-Peptid basische Aminosäurenpaare am C-Terminus und am N-Terminus hat,

22. Verfahren nach einem der Ansprüche 4 bis 21, wobei das Linker-Peptid die Aminosäuresequenz RLYRRHHRW-GRSGSPLRAHE (Sequenz-Identifizierungsnr. 1) hat.

23. Verfahren zur Herstellung eines gewünschten Polypeptids, welches die Schritte umfaßt:

(1) Transformieren von Escherichia coli-Wirtszellen mit einem Expressionsvektor, der ein Gen umfaßt, das ein Fusionsprotein codiert, das mindestens ein schützendes Polypeptid, ein gewünschtes Polypeptid und ein Linker-Peptid umfaßt, wobei das schützende Polypeptid ein Polypeptid, das von E. coli β-Galactosidase stammt, und/oder ein Polypeptid ist, das von Aminoglycosid-3'-Phosphotransferase von Transposon 903 stammt, das gewünschte Polypeptid ein Derivat von Staphylococcus aureus V8-Protease ist, das Linker-Peptid zwischen dem schützenden Polypeptid und dem gewünschten Polypeptid eine Substratstelle hat, die von einer für die Wirtszellen intrinsischen Protease spezifisch erkannt wird;
(2) Ausprägen des Gens in E. coli-Wirtszellen, um das Derivat von Staphylococcus aureus V8-Protease als inaktives Fusionsprotein zu produzieren;
(3) Aufbrechen der Zellen, um das Fusionsprotein abzutrennen und eine Fraktion zu erhalten, die E. coli ompT-Protease, die eine endogene Protease der Wirtszellen ist, und das Fusionsprotein enthält;
(4) Lösen des Fusionsproteins mit einem Denaturierungsmittel und
(5) Verringern der Konzentration des Denaturierungsmittels auf einen Wert, bei dem E. coli ompT-Protease seine Aktivität zeigt, das Linker-Peptid mit der Protease abzuspalten, so daß das gewünschte Polypeptid aus dem Fusionsprotein erhalten wird.

Revendications

1. Un procédé de production d'un polypeptide désiré comprenant les étapes consistant à :

- transformer des cellules-hôtes avec un vecteur d'expression comprenant un gène codant pour une protéine de fusion comprenant le polypeptide désiré et un polypeptide de protection ;
- cultiver les cellules-hôtes transformées de manière à exprimer ledit gène pour produire une protéine de fusion qui est insoluble et déposée intracellulairement sous la forme de corps d'inclusion ;
- rompre lesdites cellules-hôtes pour séparer lesdits corps d'inclusion ; et
- utiliser une protéase pour exciser le polypeptide désiré contenu dans la protéine de fusion et produire le polypeptide désiré ;

caractérisé en ce que la protéase utilisée pour exciser le polypeptide désiré contenu dans la protéine de fusion est une protéase endogène provenant des cellules-hôtes rompues elles-mêmes, qui a été séparée avec les corps d'inclusion.

2. Un procédé selon la revendication 1 comprenant les étapes consistant à solubiliser la protéine de fusion avec un agent dénaturant et à cliver le peptide de liaison en abaissant la concentration de l'agent dénaturant à une concentration plus basse à laquelle la protéase endogène aux cellules-hôtes manifeste une activité enzymatique, de manière à obtenir le polypeptide désiré à partir de la protéine de fusion.

3. Un procédé selon la revendication 2, dans lequel le peptide désiré est ré-enroulé en une conformation active après clivage de la protéine de fusion par la protéase en abaissant la concentration de l'agent dénaturant.

4. Un procédé selon l'une quelconque des revendications précédentes, dans lequel la protéine de fusion est représentée par la formule (1) A-L-B ou (2) A-L-B-L-C, où A et C sont des polypeptides de protection, B est un polypeptide désiré et L est un peptide de liaison contenant un site de substrat reconnu par une protéase intrinsèque aux cellules-hôtes, et la protéine de fusion est clivée dans la région du peptide de liaison L de manière à obtenir le polypeptide B désiré.

5. Un procédé selon l'une quelconque des revendications précédentes, dans lequel les cellules-hôtes sont E. coli.

6. Un procédé selon l'une quelconque des revendications précédentes, dans lequel ledit peptide désiré est physiologiquement actif et choisi dans le groupe constitué par la motiline, le glucagon, la corticotrophine (ACTH), l'hormone libératrice de la corticotrophine (CRH), la sécrétine, la somatotrophine, l'insuline, la stomatostatine (GRH), la vasopressine, l'oxytocine, la gastrine, un peptide apparenté au glucagon (GLP-1, GLP-2, 7-36 amides), la cholécystokinine, un polypeptide intestinal vaso-actif (VIP), un polypeptide activant l'adénolate cyclase hypophysaire (p.a.c.a.p.), l'hormone de libération de la gastrine, la galanine, la thyrotrophine (TSH), l'hormone de libération de la lutéinostimuline (LH-RH), la calcitonine, la parathormone (PTH, PTH(1-34), PTH(1-84), un peptide histidine-isoleucine (PHI), un neuropeptide Y (nP.Y)), un peptide YY (P.YY), un polypeptide pancréatique (P.P.), la somatostatine, TGF-α, TGF-β, le facteur de croissance nerveuse, le facteur de croissance fibroblastique, la relaxine, la prolactine, un peptide natriurétique, l'angiotensine, et le facteur nourricier dérivé du cerveau.

7. Un procédé selon la revendication 6, dans lequel le peptide natriurétique est choisi dans le groupe constitué par ANP, BNP et CNP.

8. Un procédé selon l'une quelconque des revendications 1 à 5, dans lequel le polypeptide désiré est une enzyme.

9. Un procédé selon la revendication 8, dans lequel le polypeptide désiré est une enzyme protéolytique.

10. Un procédé selon la revendication 9, dans lequel l'enzyme protéolytique est choisie dans le groupe constitué par l'endopeptidase KEX2, la protéase V8 de Staphylococcus aureus et un dérivé de la protéase V8 de Staphylococcus aureus.

11. Un procédé selon l'une quelconque des revendications précédentes, dans lequel ledit polypeptide désiré a de 20 à 800 résidus d'acides aminés.

12. Un procédé selon la revendication 10, dans lequel le dérivé de protéase V8 de S. aureus a la séquence en acides aminés indiquée par le soulignement sur la figure 4, ou la figure 6, ou une séquence en acides aminés telle qu'illustrée sur la figure 14.

13. Un procédé selon l'une quelconque des revendications 2 à 12, dans lequel l'agent dénaturant est choisi dans le groupe consistant en l'urée, le chlorhydrate de guanidine et des tensioactifs.

14. Un procédé selon la revendication 13, dans lequel l'agent dénaturant est l'urée 1 à 8 M.

15. Un procédé selon la revendication 14, dans lequel la concentration de l'urée pendant le clivage de la protéine de fusion est de 1 à 6 M.

16. Un procédé selon l'une quelconque des revendications précédentes, dans lequel le polypeptide de protection est

dérivé de β-gal de E. coli et/ou de l'aminoglycoside 3'-phosphotransférase de transposon 903.

17. Un procédé selon l'une quelconque des revendications 4 à 16, dans lequel le polypeptide de protection A est dérivé de β-gal de E. coli et le polypeptide de protection C est dérivé de l'aminoglycoside 3'-phosphotransférase de transposon 903.

18. Un procédé selon l'une quelconque des revendications précédentes, dans lequel la protéase est la protéase ompT de E. coli.

19. Un procédé selon l'une quelconque des revendications précédentes, dans lequel la protéase est ompT et la protéine de fusion est clivée par la protéase dans une solution contenant environ 4M d'urée.

20. Un procédé selon l'une quelconque des revendications 4 à 19, dans lequel le peptide de liaison contient de 2 à 50 acides aminés et une ou deux paires d'acides aminés basiques.

21. Un procédé selon l'une quelconque des revendications 4 à 20, dans lequel ledit peptide de liaison a des paires d'acides aminés basiques à l'extrémité C-terminale et à l'extrémité N-terminale.

22. Un procédé selon l'une quelconque des revendications 4 à 21, dans lequel le peptide de liaison a la séquence en acides aminés :

RLYRRHHRWGRSGSPLRAHE (SEQ ID N° :1).

23. Un procédé de production d'un polypeptide désiré, comprenant les étapes consistant à :

(1) transformer des cellules-hôtes Escherichia coli par un vecteur d'expression comprenant un gène codant pour une protéine de fusion comprenant au moins un polypeptide de protection, un polypeptide désiré et un peptide de liaison, dans lequel le polypeptide de protection est un polypeptide dérivé de la β-galactosidase de E. coli et/ou un polypeptide dérivé d'une aminoglycoside-3'-phosphotransférase de transposon 903, le polypeptide désiré est un dérivé de la protéase V8 de Staphylococcus aureus, le peptide de liaison entre ledit polypeptide de protection et ledit polypeptide désiré contient un site de substrat spécifiquement reconnu par une protéase intrinsèque aux cellules-hôtes ;
(2) exprimer ledit gène dans des cellules-hôtes E. coli pour produire le dérivé de protéase V8 de Staphycoccus aureus sous la forme d'une protéine de fusion inactive;
(3) rompre les cellules de manière à séparer la protéine de fusion, et obtenir une fraction contenant la protéase ompT de E. coli qui est une protéase endogène des cellules et la protéine de fusion ;
(4) solubiliser la protéine de fusion avec un agent dénaturant; et
(5) abaisser la concentration de l'agent dénaturant jusqu'à un niveau auquel la protéase ompT de E. coli manifeste son activité pour cliver le peptide de liaison avec la protéase de manière à obtenir le polypeptide désiré contenu dans la protéine de fusion.

Fig.1(a)

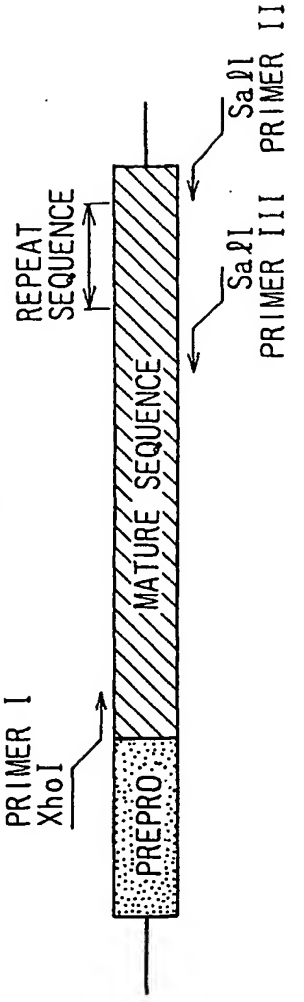


Fig.1(b)

PRIMER I ;5' ACCGCTCGAGGTTATATTACCAAATAACGAT3' (SEQ ID NO:2)
XhoI

PRIMER II ;5' CTTAATGTCGACTTAAGCTGCATCTGGATT3, (SEQ ID NO:3)
SalI

PRIMER III;5' TCGCGTCGACTTATTGGTCATCGTTGGCAAA3, (SEQ ID NO:4)
SalI

Fig.2

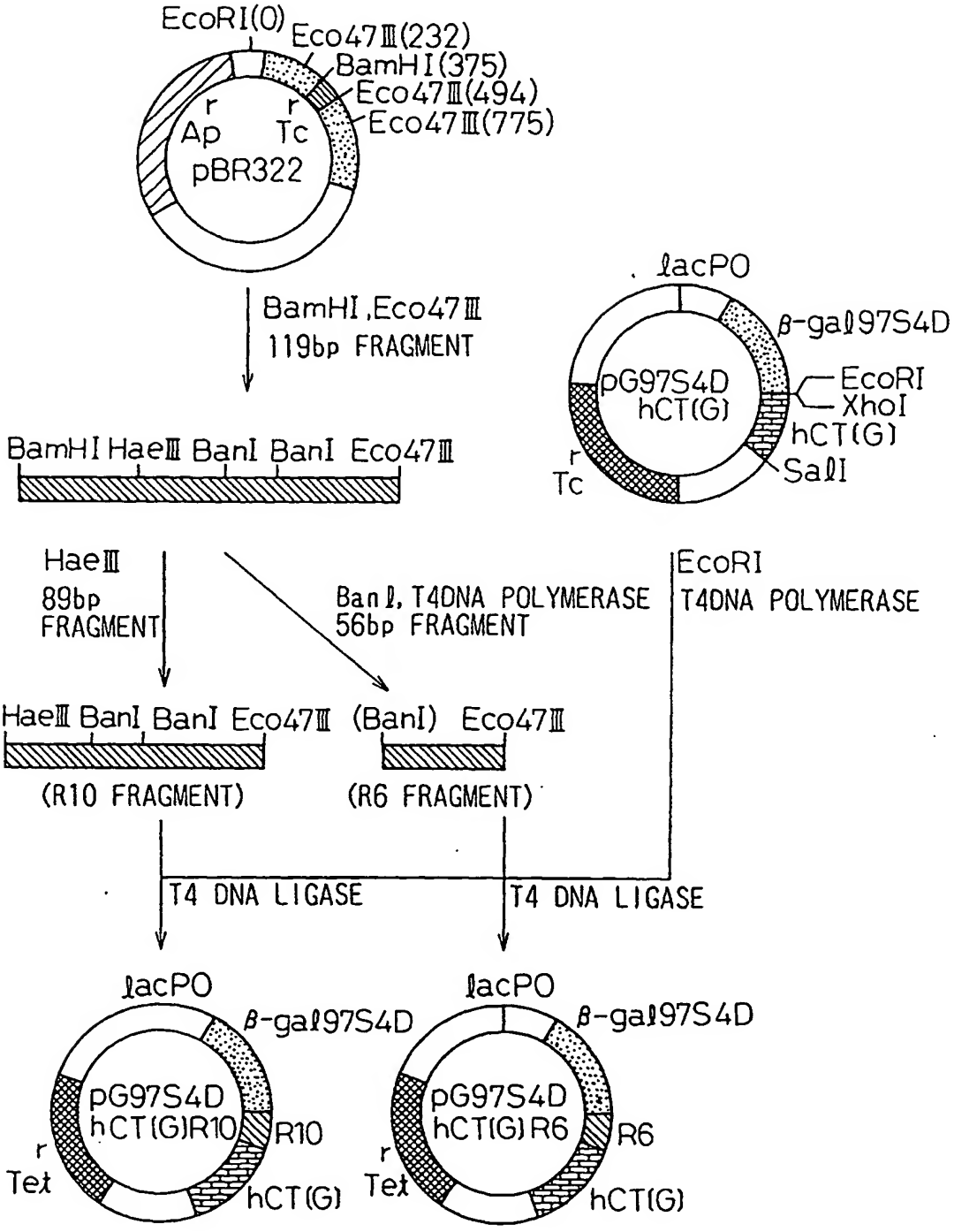

EcoRI(0)
Eco47Ⅲ(232)
BamHI(375)
Eco47Ⅲ(494)
Eco47Ⅲ(775)
r
Ap
r
Tc
pBR322
BamHI,Eco47Ⅲ
119bp FRAGMENT
BamHI HaeⅢ BanI BanI Eco47Ⅲ
lacPO
β-gal97S4D
pG97S4D
hCT(G)
EcoRI
XhoI
hCT(G)
SalI
r
Tc
HaeⅢ
89bp
FRAGMENT
BanⅠ, T4DNA POLYMERASE
56bp FRAGMENT
EcoRI
T4DNA POLYMERASE
HaeⅢ BanI BanI Eco47Ⅲ
(R10 FRAGMENT)
(BanI) Eco47Ⅲ
(R6 FRAGMENT)
T4 DNA LIGASE
T4 DNA LIGASE
lacPO
β-gal97S4D
pG97S4D
hCT(G)R10
R10
r
Tet
hCT(G)
lacPO
β-gal97S4D
pG97S4D
hCT(G)R6
R6
r
Tet
hCT(G)

Fig.3

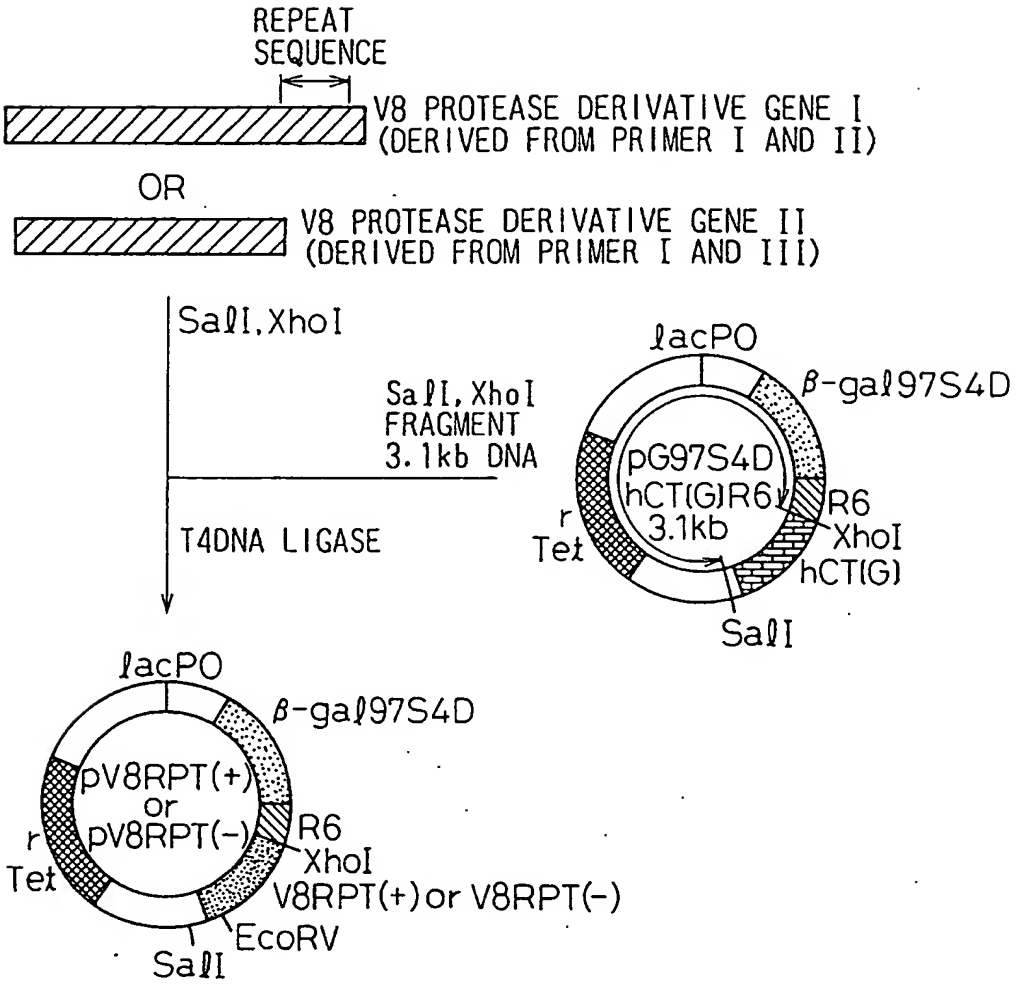

REPEAT
SEQUENCE
V8 PROTEASE DERIVATIVE GENE I
(DERIVED FROM PRIMER I AND II)
OR
V8 PROTEASE DERIVATIVE GENE II
(DERIVED FROM PRIMER I AND III)
SalI, XhoI
SalI, XhoI
FRAGMENT
3.1kb DNA
T4DNA LIGASE
lacPO
β-gal97S4D
pG97S4D
hCT(G)R6
3.1kb
r
Tet
R6
XhoI
hCT(G)
SalI
lacPO
β-gal97S4D
pV8RPT(+)
or
pV8RPT(-)
r
Tet
R6
XhoI
V8RPT(+) or V8RPT(-)
EcoRV
SalI

Fig.4(a)

1 50
MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSDDARTDRPSQ
100
QLRSLNGEWRFAWFPAPEAVPDSLLDSDLPEADTVVVPSNWQMHGYDAEL
150
RLYRRHHRWGRSGSPLRAHEQFLEVILPNNDRHQITDTTNGHYAPVTYIQ
200
VEAPTGTFIASGVVVGKDTLLTNKHVVDATHGDPHALKAFPSAINQDNYP
250
NGGFTAENITKYSGEGDLAIVKFSPNEQNKHIGEVVKPATMSNNAETQVN
300
QNITVTGYPGDKPVATMWESKGKITYLKGEAMQYDLSTTGGNSGSPVFNE
344
KNEVIGIHWGGVPNEFNGAVFINENVRNFLKQNIEDIHFANDDQ (SEQ ID NO:5)

AMINO ACID SEQUENCE OF FUSION PROTEIN ENCODED IN pV8RPT (-)
(V8 PROTEASE REGION IS UNDERLINED)

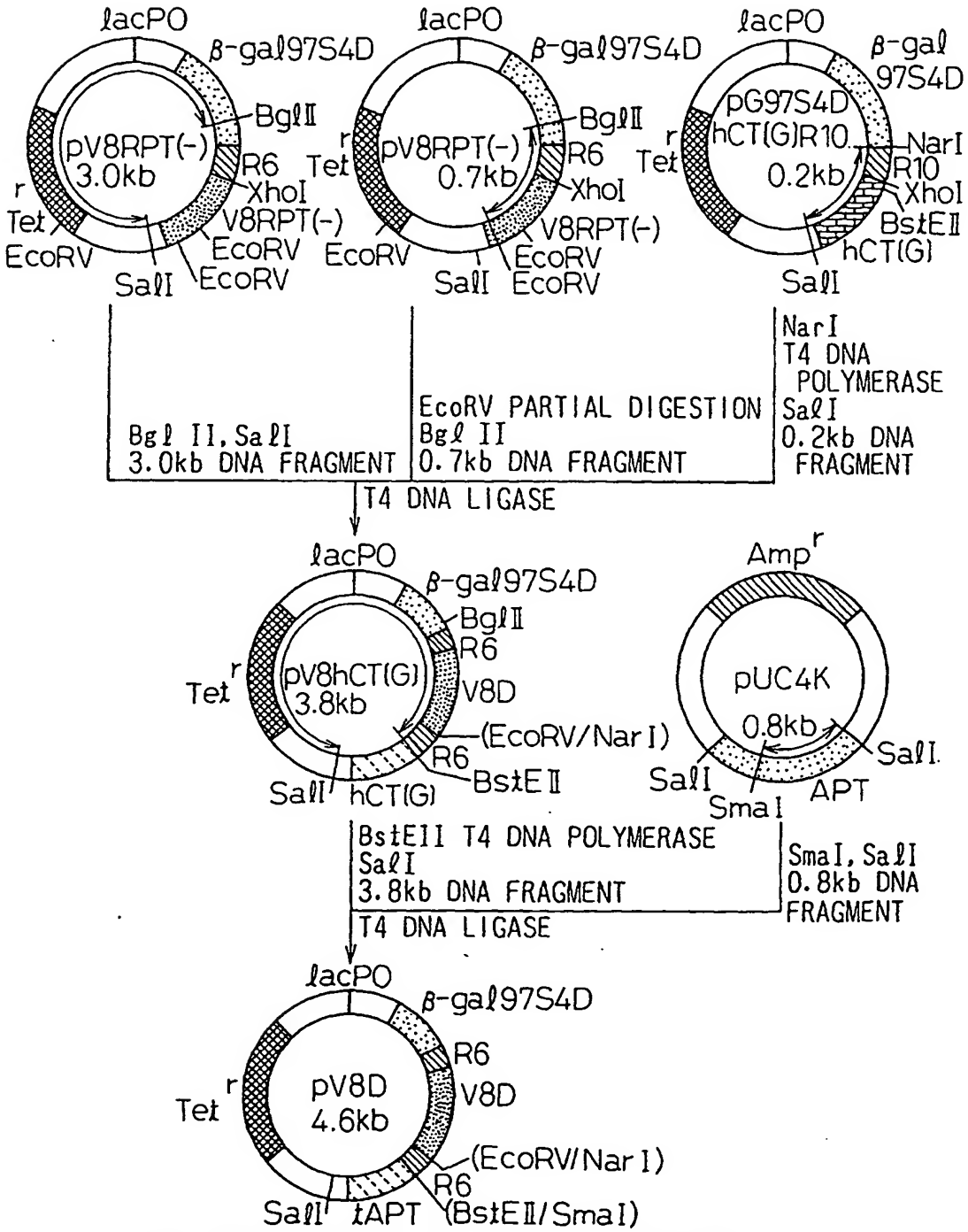
Fig.4(b)

1
MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSDDARTDRPSQ 50
QLRSLNGEWRFAWFPAPEAVPDSLLDSDLPEADTVVVPSNWQMHGYDAEL 100
RLYRRHHRWGRSGSPLRAHEQFLEVILPNNDRHQITDTTNGHYAPVTYIQ 150
VEAPTGTFIASGVVVGKDTLLTNKHVVDATHGDPHALKAFPSAINQDNYP 200
NGGFTAENITKYSGEGDLAIVKFSPNEQNKHIGEVVKPATMSNNAETQVN 250
QNITVTGYPGDKPVATMWESKGKITYLKGEAMQYDLSTTGGNSGSPVFNE 300
KNEVIGIHWGGVPNEFNGAVFINENVRNFLKQNIEDIHFANDDQPNNPDN 350
PDNPNNPDNPNNPDEPNNPDNPNNPDNPDNGDNNNSDNPDAA 392 (SEQ ID NO:6)

AMINO ACID SEQUENCE OF FUSION PROTEIN ENCODED IN pV8RPT (+)
(V8 PROTEASE REGION IS UNDERLINED)

EP 0 700 995 B1

Fig. 5



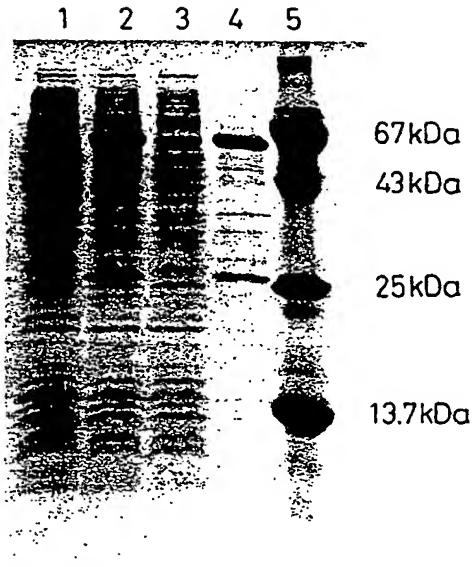
APT; AMINOGLYCOSIDE 3'-PHOSPHOTRANSFERASE
tAPT; TRUNCATED AMINOGLYCOSIDE 3'-PHOSPHOTRANSFERASE

Fig.6

```
1                                                                  50
MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSDDARTDRPSQ
                                                                  100
QLRSLNGEWRFAWFPAPEAVPDSLLDSDLPEADTVVVPSNWQMHGYDAEL
                  ↓                                               150
RLYRRHHRWGRSGSPLRAHEQFLEVILPNNDRHQITDTTNGHYAPVTYIQ
                                                                  200
VEAPTGTFIASGVVVGKDTLLTNKHVVDATHGDPHALKAFPSAINQDNYP
                                                                  250
NGGFTAENITKYSGEGDLAIVKFSPNEQNKHIGEVVKPATMSNNAETQVN
                                                                  300
QNITVTGYPGDKPVATMWESKGKITYLKGEAMQYDLSTTGGNSGSPVFNE
                                                  ↓               350
KNEVIGIHWGGVPNEFNGAVFINENVRNFLKQNIEDRLYRRHHRWGRSGS
                                                                  400
PLRAHEQFLECGNGKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNC
                                                                  450
PFNSDRVFRLAQAQSRMNNGLVDASDFDDERNGWPVEQVWKEMHKLLPFS
                                                                  500
PDSVVTHGDFSLDNLIFDEGKLIGCIDVGRVGIADRYQDLAILWNCLGEF
                                                  532
SPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF (SEQ ID NO:7)
```

AMINO ACID SEQUENCE OF FUSION PROTEIN ENCODED IN pV8D
V8 PROTEASE REGION IS UNDERLINED
↓ CLEAVAGE SITE FOR OmpT PROTEASE

Fig.7



1. TOTAL PROTEIN PRIOR INDUCTION
2. TOTAL PROTEIN AFTER INDUCTION
3. PROTEIN IN SOLUBLE FRACTION AFTER INDUCTION
4. PROTEIN IN INSOLUBLE FRACTION AFTER INDUCTION
5. MOLECULAR WEIGHT MARKER

Fig.8A

INCLUSION BODIES DERIVED FROM E.coli JM101

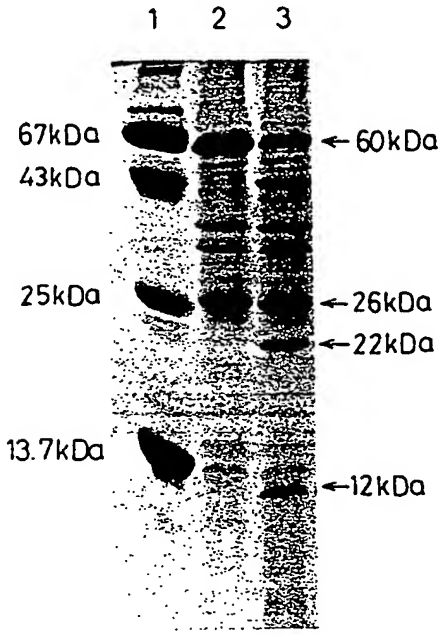
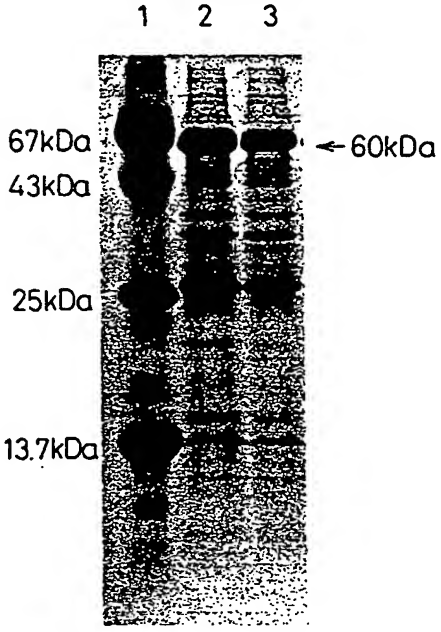


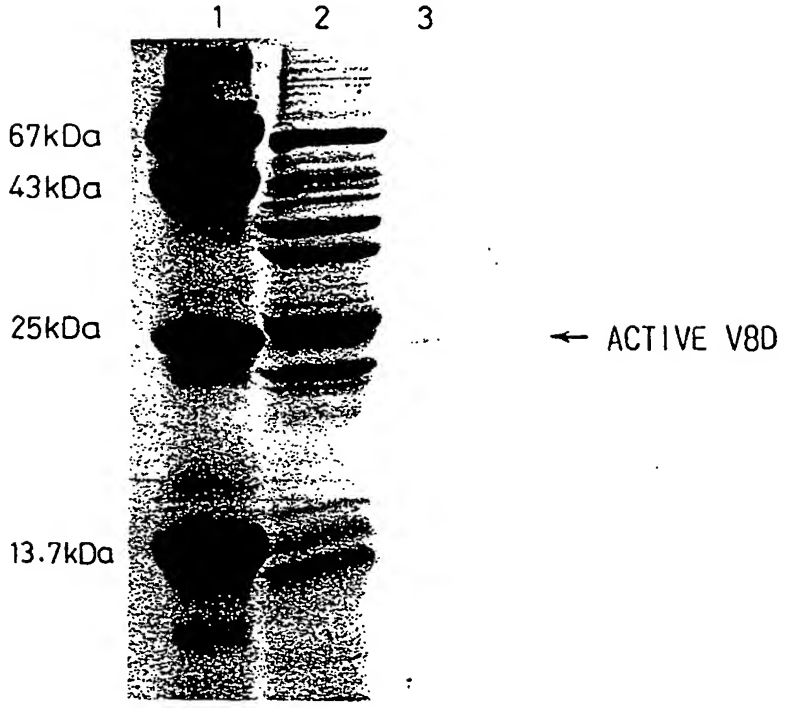
Fig.8B

INCLUSION BODIES DERIVED FROM E.coli W3110M25 (ompT)



PROCESSING OF V8D FUSION PROTEIN BY ompT PROTEASE
1. MOLECULAR WEIGHT MARKER
2. IMEDIATEHS AFTER SOLUBILIZATION WITH UREA
3. AFTER 2 HOURS

Fig.9



1. MOLECULAR WEIGHT MARKER
2. AFTER ompT PROTEASE TREATMENT
3. AFTER REFOLDING

Fig.10A

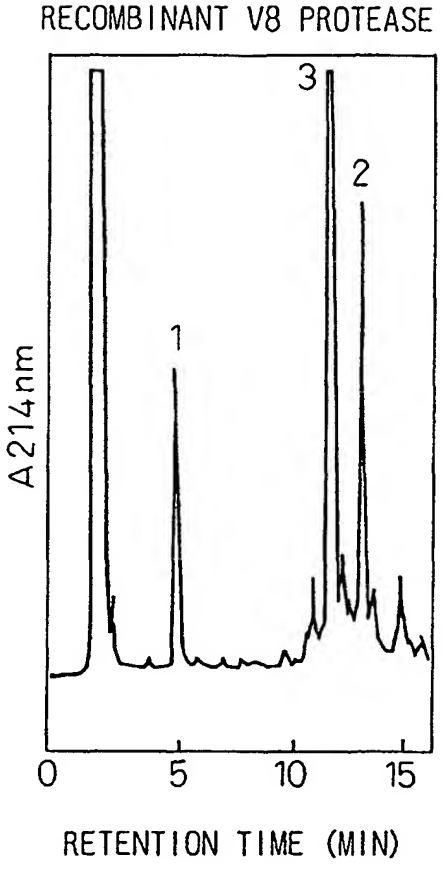
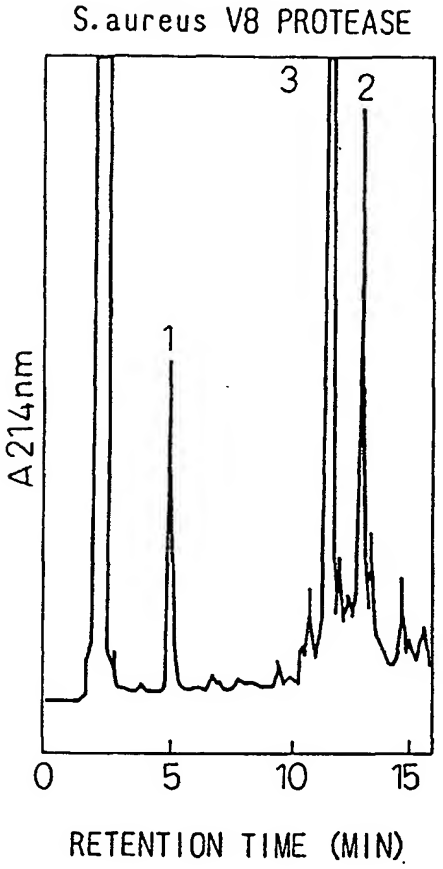


Fig.10B



PEAK1. hCT[G]
PEAK2. FUSION PROTEIN OF HUMAN CALCITONIN PRECURSOR (hCT[G])
PEAK3. β-GALACTOSIDASE DERIVATIVE

Fig.11

PRIMER a; 5' ATCGTTGGCCATATGGATATCTTCAATATT (SEQ ID NO:8)
NdeI

PRIMER b; 5' GACTTATTGGTCATCGAGCTCAAAATGGATATC (SEQ ID NO:9)
SacI

PRIMER c; 5' GACTTATTGGTCGAGCTCGGCAAAATGGAT (SEQ ID NO:10)
SacI

PRIMER d; 5' ATCTGGGTTGAGCTCATCGTTGGCAAAATGGAT (SEQ ID NO:11)
SacI

PRIMER e; 5' ATCTGGTTGGAGCTCTTGGTCATCGTTGGCAAA (SEQ ID NO:12)
SacI

PRIMER f; 5' ACAAAATCATATGGAACGCCTATATCGCCGACAT (SEQ ID NO:13)
NdeI

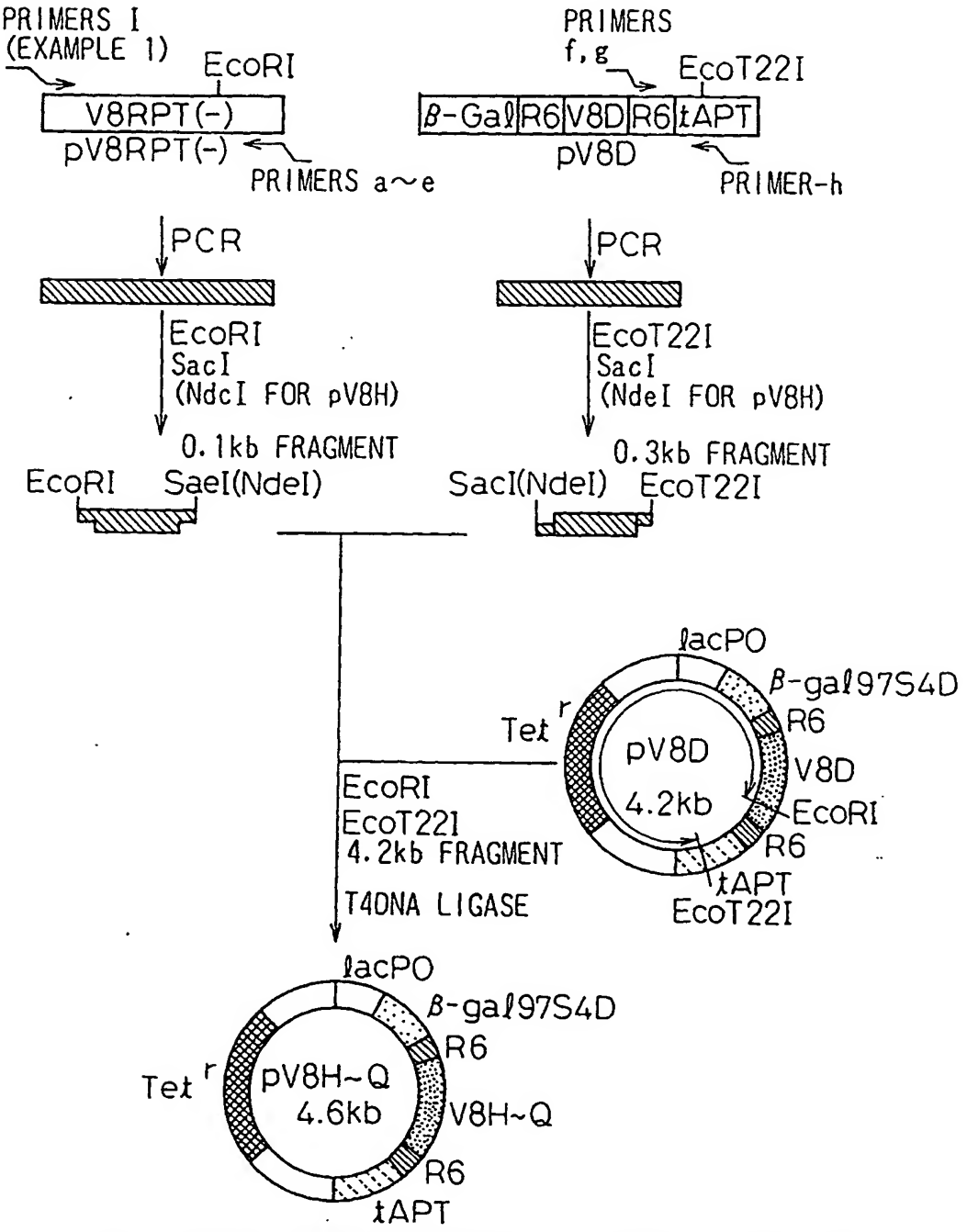
PRIMER g; 5' AATATTGAAGAGCTCCGCCTATATCGCCGACAT (SEQ ID NO:14)
SacI

PRIMER h; 5' GAATGGCAAAAGCTTATGCATTTCTTT (SEQ ID NO:15)
EcoT22I

PRIMERS a~e: CORRESPONDING TO V8 PROTEASE GENE
PRIMERS f, g: CORRESPONDING TO R6 SEQUENCE GENE
PRIMER h: CORRESPONDING TO AMINOGLUCOSIDE 3'-PHOSPHOTRANSFERSE

EP 0 700 995 B1

Fig.12



tAPT; TRUNCATED AMINOGLYCOSIDE 3'-PHOSPHOTRANSFERAS

Fig.13

PLASMID	C-TERMINAL SEQUENCE OF V8 PROTEASE REGION (209)	INCLUSION BODIES FORMATION	
pV8D	---NIED-RLYRRHHR -----	+	(SEQ ID NO:16)
pV8H	---NIEDIH-MERLYRRHHR -----	+	(SEQ ID NO:17)
pV8F	---NIEDIHF-ELRLYRRHHR -----	+	(SEQ ID NO:18)
pV8A	---NIEDIHFA-ELRLYRRHHR -----	−	(SEQ ID NO:19)
pV8D2	---NIEDIHFAND-ELRLYRRHHR -----	−	(SEQ ID NO:20)
pV8Q	---NIEDIHFANDDQ-ELRLYRRHHR -----	−	(SEQ ID NO:21)

+ −PRESENCE OR ABSENCE OF INCLUSION BODIES

□ AMINO ACID SEQUENCE DERIVED FROM V8 PROTEASE

(DESCRIBED FROM Asn (N) AT 209 POSITION FROM N-TERMINAL OF MATURE V8 PROTEASE)

EP 0 700 995 B1

Fig.14(a)

```
                                                                  50
VILPNNDRHQITDTTNGHYAPVTYIQVEAPTGTFIASGVVVGKDTLLTNK
                                                                 100
HVVDATHGDPHALKAFPSAINQDNYPNGGFTAENITKYSGEGDLAIVKFS
                                                                 150
PNEQNKHIGEVVKPATMSNNAETQVNQNITVTGYPGDKPVATMWESKGKI
                                                                 200
TYLKGEAMQYDLSTTGGNSGSPVFNEKNEVIGIHWGGVPNEFNGAVFINE
                213
NVRNFLKQNIEDI (SEQ ID NO:22)
```

Fig.14(b)

```
                                                                  50
VILPNNDRHQITDTTNGHYAPVTYIQVEAPTGTFIASGVVVGKDTLLTNK
                                                                 100
HVVDATHGDPHALKAFPSAINQDNYPNGGFTAENITKYSGEGDLAIVKFS
                                                                 150
PNEQNKHIGEVVKPATMSNNAETQVNQNITVTGYPGDKPVATMWESKGKI
                                                                 200
TYLKGEAMQYDLSTTGGNSGSPVFNEKNEVIGIHWGGVPNEFNGAVFINE
                214
NVRNFLKQNIEDIH (SEQ ID NO:23)
```


EP 0 700 995 B1

Fig.14(c)

```
                                                                  50
VILPNNDRHQITDTTNGHYAPVTYIQVEAPTGTFIASGVVVGKDTLLTNK
                                                                 100
HVVDATHGDPHALKAFPSAINQDNYPNGGFTAENITKYSGEGDLAIVKFS
                                                                 150
PNEQNKHIGEVVKPATMSNNAETQVNQNITVTGYPGDKPVATMWESKGKI
                                                                 200
TYLKGEAMQYDLSTTGGNSGSPVFNEKNEVIGIHWGGVPNEFNGAVFINE
                  215
NVRNFLKQNIEDIHF (SEQ ID NO:24)
```